MULTIMEDIA COMMUNICATIONS TECHNICAL COMMITTEE
IEEE COMMUNICATIONS SOCIETY
http://www.comsoc.org/~mmc

E-LETTER

Vol. 6, No. 2, February 2011

IEEE COMMUNICATIONS SOCIETY

CONTENTS

SPECIAL ISSUE ON VIDEO-AWARE WIRELESS NETWORKS

Video-Aware Wireless Networks

Guest Editor: Thinh Nguyen, Oregon State University,USA
thinhq@eecs.oregonstate.edu

Advances in wireless communication technologies have contributed to an explosive growth of video applications over wireless networks in recent years. This E-letter includes five papers. Each provides a glimpse at the challenges and opportunities in designing efficient video transmission systems over wireless networks. The paper topics range from developing novel wireless video applications and systems to exploring new video-aware protocols and distortion metrics.

In the first paper, titled "Video-Aware Cross-Layer Design" by Li Zhu, F. Richard Yu, and Bin Ning, the authors investigate the optimal policy for hand-off in an 802.11p-based train-ground video communication networks used to deliver real-time multimedia information in a metro Passenger Information System (PIS). The optimization problem is formulated as a stochastic semi-Markov Decision Process. Simulation results show substantial improvement in video quality over two other heuristic policies.

In the second paper "Video Streaming in FiWi Access Networks" by Navid Ghazisaidi, Martin Maier, and Martin Reisslein, the authors explore advanced MAC protocols for video delivery over FiWi networks that employ both radio-over-fiber (RoF) and radio-and-fiber(R&F) technologies. They provide the first study of efficient integrated MAC mechanisms for FiWi networks with streaming prerecorded video traffic in downstream direction. Their preliminary results indicate that video frame prefetching with channel probing reduces the playback starvation probabilities significantly.

The third paper "Prediction of Transmission Distortion for Wireless Video Communication" by Zhifeng Chen and Dapeng Oliver Wu addresses a number of issues involving transmission distortion prediction which are useful for optimally selecting video encoding mode, controlling the cross-layer encoding rate, and scheduling packets for wireless transmissions. Their transmission distortion prediction algorithm shows a 1.44dB over the well-known ROPE algorithm under packet error probability of 1%.

The fourth paper "MSE Cross-Layer Optimization Criteria: What Else?" by Laura Toni and Pascal Frossard investigates the advantages and limitations of the mathematical metrics and visual perception metrics used in video quality assessment. The authors conclude that the current mathematical metrics such as MSE, is simple to evaluate, models the visual fidelity reasonably well, but only provides a coarse approximation of the actual visual distortion. On the other hand, perceptual metrics are suitable to benchmark systems, but their use in cross-layer design is not trivial.

The last paper "Video-on-Demand Broadcasting to Mobile Wireless Users" by Duc Tran is a short survey on techniques and architecture for video-on-demand applications in mobile wireless networks. The author briefly discusses a variety of approaches to address the issues of limited bandwidth and coverage of wireless networks. Specifically, a number of efficient broadcast protocols and caching techniques are proposed. The paper concludes by pointing out the challenges as well as tremendous benefits of future mobile VOD systems.

Thinh Nguyen is an Assistant Professor at the School of Electrical Engineering and Computer Science of the Oregon State University. He received his M.S and Ph.D. from the University of California, Berkeley. He has served on many technical program committees. He is an associate editor of the IEEE Transactions on Multimedia and the Peer-to-Peer Networking and Applications. His research interests include Multimedia Networking and Processing, Wireless Networks, and Network Coding.

Video-aware Cross-layer Design in Train-ground Wireless Communication Networks

Li Zhu, Beijing Jiaotong University, P.R. China
F. Richard Yu, Carleton University, Ottawa, ON, Canada
Bin Ning, Beijing Jiaotong University, P.R. China

1. Introduction

Recent population explosion causes severe traffic jam problems in big cities. More and more cities have been devoted to building comfortable, convenient and efficient metro systems to relieve the pressure on already busy roads. Based on computer system and multi-media network technologies, a metro Passenger Information System (PIS) provides passengers with real-time dynamic multimedia information, such as the train schedule, live news, live matches, finance and economics, weather forecast, etc., in stations or on trains [1]. When fire, earthquake or terrorist attack occurs, metro PISs can broadcast various kinds of dynamical emergency evacuation indications.

Train-ground wireless video communication network is one of the key subsystems in metro PISs. An advanced train-ground video communication network is able to guarantee a high quality transmission of video information between the high speed train and ground, which can help provide the passengers with high quality video programs and give them a comfortable journey when they are commuting.

In this letter, we study the video transmission optimization issues in metro PIS train-ground video communication networks based on IEEE 802.11p [2]. Fountain codes are used in our proposed train-ground video communication network as the application layer forward error correction (FEC) scheme. We take a cross-layer design approach to jointly optimize application layer parameters and handoff decisions to improve video transmission quality in metro PISs. Based on the sensed channel condition, the wireless controller on the train makes handoff decisions and adapts the application layer parameters to optimize the video transmission quality of service (QoS).

2. The Proposed Train-Ground Video Communication Network based on IEEE 802.11p and Fountain Codes

The idea of fountain codes is first proposed in [3]. Followed by some practical realization like online codes, LT codes and raptor codes, fountain codes have gained significant attention in video streaming in recent years [4]. Fountain codes are different from the physical layer FEC codes where channel encoding is performed for a fixed channel rate and all encoded packets are generated prior to transmission. The fountain codes idea is proven to be efficient for large source data sizes [5] in the case of video data. Furthermore, this idea is very suitable to metro PISs where frequent handoffs happen. The fountain codes can recover the lost packets caused by handoffs without retransmission.

Fig. 1 describes the proposed train-ground video communication network in metro PISs. Fountain codes are applied in the application layer. The coded application layer data packets, being encapsulated in RTP, are then transported over the proposed network. The IP packets are encapsulated by the IEEE 802.11p MAC and physical layers.

A critical issue in the above system is the handoff decision policy, i.e., when to perform handoff, and the corresponding application layer parameters adaptation policy. In order to optimize identify the optimal policies, we formulate the handoff decision and application layer parameters adaptation problem as a stochastic semi-Markov Decision Process (SMDP).

An SMDP model consists of the following five elements: (1) decision epochs, (2) states, (3) actions, (4) rewards, and (5) transition probabilities, which will be described in the following.

1) Decision Epoch, Action and State: In our SMDP model, the wireless controller on the train has to make a decision whenever a certain time period has elapsed. The instant times are called *epochs.* At each decision epoch, the wireless controller on the train first has to decide whether the connection should use the current chosen AP or connect to the next AP (we assume the SA on the train won't be in the coverage of 3 successive APs). The application layer parameters should be decided as well. The current composite action $a(t) \in A$ is denoted by

$$a(t) = \{a_h(t),\ a_\beta(t), a_\rho(t)\}, \qquad (1)$$

where $a_h(t)$ is the handoff decision $a_h(t) = 1$means handoff to the next AP; $a_h(t) = 0$ means stay in the old AP), $a_\beta(t)$ is the intra updating rate decision(0< $a_\beta(t)$ <1),and $a_\rho(t)$ is the raptor codes parity rate decision($a_\rho(t)$>0).

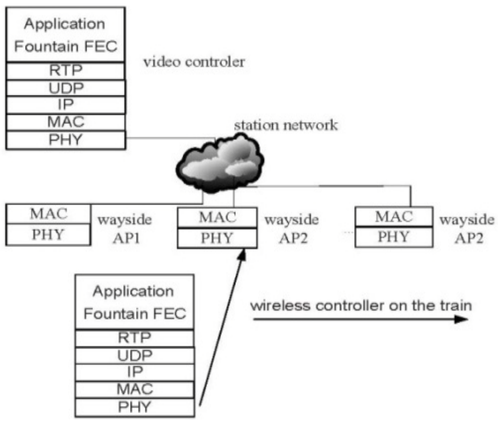


Figure 1. The proposed train-ground video communication network

The current composite state $s(t) \in S$ is given as

$$s(t) = \{\gamma_1(t), \gamma_2(t)\}, \quad (2)$$

where $\gamma_1(t)$ and $\gamma_2(t)$ are the measured signal noise ratio (SNR) from two successive APs, respectively.

2) Reward Function: In our formulation, we define the reward function as

$$r(s(t), a(t)) = \frac{1}{D(s(t), a(t))}, \quad (3)$$

where $D(s(t), a(t))$ reflects the end-to-end total distortion under the current state and chosen action. When $a_h(t) = 0$, which means a handoff does not occur, the distortion is mainly caused by IEEE 802.11p link layer packet losses rate p_l and the known parity rate decision $a_\rho(t)$. The final end-to-end total distortion $D(s(t), a(t))$ can be obtained from [6].

We assume the packet loss rate in IEEE 802.11p link layer is constant when a handoff occurs. We denote it as p_h. Then the end-to-end total distortion is derived just the same as the no handoff situation.

3) State Transition Probability: Given the current state $s(t) = \{\gamma_1(t), \gamma_2(t)\}$ and the chosen action $a(t)$, the probability function of the next state $s(t+1) = [\gamma_1(t+1), \gamma_2(t+1)]$ is given by

$$P(s(t+1)|s(t), a(t)) = P[\gamma_1(t+1)|\gamma_1(t)] * P[\gamma_2(t+1)|\gamma_2(t), \quad (4)$$

where $P[\gamma_1(t+1)|\gamma_1(t)$ and $P[\gamma_2(t+1)|\gamma_2(t)]$ are the channel state transition probabilities for the two wireless links, respectively.

There are various algorithms available to solve the SMDP problem. We use value iteration algorithm (VIA) in this letter to determine a stationary deterministic optimal policy.

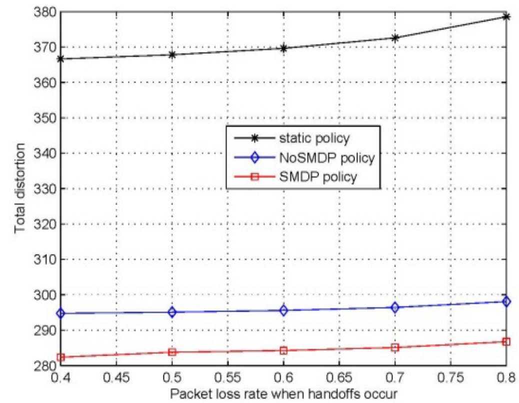


Figure 2. The total distortion under three policies

3. Simulation Results

We compare the performance of our proposed policies with two other heuristic policies. For the first heuristic policy, the train wireless controller decides to handoff whenever the other AP provides better QoS, and the raptor coding parity rate and intra updating rate are both fixed. We denote this policy as static policy. For the second heuristic policy, we adapt the raptor coding parity rate and intra updating rate according to the channel state, but the wireless controller makes handoff decisions only based on current channel state. We denote this policy as NoSMDP policy.

Fig.2 shows the end-to-end total distortion for the three policies. The SMDP policy gives the lowest distortion for all different p_h, and it significantly decrease the total distortion compared to the other two policies. This is because both of the two heuristic policies make their handoff decisions by their current channel information, and neither of the heuristic policies considers the dynamic transition of the wireless channel in metro PISs, which is very important information to make handoff decisions.

4. Conclusions

Providing messengers with real-time multimedia information is one of the most important applications in metro Passenger Information Systems (PISs). Frequent train handoffs can cause significant video distortion in metro PISs. In this paper, we presented an integrated approach to jointly optimize application layer parameters and handoff decisions to optimize video transmissions over PISs. We have proposed a metro PIS train-ground video communication network based on fountain codes and IEEE 802.11p. The video transmission problem was modeled as a semi-Markov Decision Process (SMDP). Simulation results were presented to show that the proposed SMDP based optimization algorithm can significantly decrease the video transmission distortion.

References

[1] K. G. Zografos, K. N. Androutsopoulos, and V. Spitadakis, “Design and assessment of an online passenger information system for integrated multimodal trip planning,” IEEE Trans. Intell. Transp. Syst., vol. 10,pp. 311 – 323, Jun. 2009.

[2] IEEE802.11p/D3.0, “Draft amendment for wireless access in vehicular environments (WAVE),” July 2007.

[3] J. W. Byers, M. Luby, M. Mitzenmacher, and A. Rege, “A digital fountain approach to reliable distribution of bulk data,” in Proc. ACM SIGCOMM’98, pp. 56–67, 1998.

[4] J.-P. Wagner, J. Chakareski, and P. Frossard, “Streaming of scalable video from multiple servers using rateless codes,” in Proc. IEEE Int’l Conf. on Multimedia and Expo, pp. 1501–1504, 2006.

[5] M. Luby, “Lt codes,” in Proc. IEEE Symp. Foundations of Computer Science, pp. 271 – 280, 2002.

[6] Z. He, J. Cai, and C. W. Chen, “Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding,” IEEE Trans. Circ. Syst. for Video Tech., vol. 12, no. 6, pp. 511–523, 2002.

Li Zhu is currently a PhD student at Beijing Jiaotong university and a visiting PhD student at Carleton University, Ottawa, ON, Canada. His research interests include Communication Base Train Ground Communication (CBTC) systems, cross layer design in train ground communication systems.

F. Richard Yu is currently an Assistant Professor at Carleton University. His research interests include cross-layer design, security and QoS provisioning in wireless networks. Dr. Yu is a senior member of the IEEE. He serves on the editorial boards of several journals. For more details, see http://www.csit.carleton.ca/~fyu/

Bin Ning is currently a professor at Beijing Jiaotong University. His research mainly focus on High speed train control system and railway transportation train control system, including Main Locomotive Signal, communication based train control system (CBTC), intelligent transportation, fault-tolerant design of signal system, fault diagnosis, system reliability and security design.

Video Streaming in FiWi Access Networks

Navid Ghazisaidi, Arizona State University, Tempe, AZ, USA
navid@asu.edu
Martin Maier , Optical Zeitgeist Laboratory, INRS, Montreal, QC, Canada
maier@emt.inrs.ca
Martin Reisslein, Arizona State University, Tempe, AZ, USA
reisslein@asu.edu

1. Emergence of FiWi networks

Bimodal fiber-wireless (FiWi) networks may be considered the endgame of broadband access. FiWi access networks aim at seamlessly converging the latest optical and wireless broadband technologies, capitalizing on the capacity of optical networks and the ubiquity and mobility of wireless networks. FiWi networks involve the deployment of both radio-over-fiber (RoF) and radio-and-fiber (R&F) technologies [1]. Significant progress has been made at the physical layer toward the commercial realization of RoF networks. In a recently reported RoF field trial, the Georgia Institute of Technology successfully demonstrated the delivery of 270 Mb/s standard definition (SD) and 1.485 Gb/s high definition (HD) real-time video streams using 2.4 and 60 GHz millimeter-wave transmissions over standard single-mode fiber [2].

RoF networks are well suited for medium access control (MAC) protocols that deploy centralized polling and scheduling, e.g., cellular and WiMAX networks. However, for distributed MAC protocols, such as the widely deployed DCF in IEEE 802.11 WLANs, the additional fiber propagation delay may exceed certain MAC protocol timeouts, e.g., ACK timeout. As a consequence, optical fiber can be deployed in WLAN-based RoF networks only up to 1948 meters. R&F networks can avoid these limitations by means of protocol translation at the optical-wireless interface and controlling access to the optical and wireless media separately from each other [3].

1.1 Video delivery over FiWi Networks

Recently, the University of California Davis developed an R&F prototype by integrating Ethernet passive optical networks (EPONs) with an IEEE 802.11g WLAN-based wireless mesh network (WMN) [4]. The reported results show that the quality of video transmissions sharply deteriorates for an increasing number of wireless hops. In fact, the video client shows a blank screen after four wireless hops. These experimental results clearly show that a more involved study of integrated EPON/WLAN-based WMN networks are needed to support video traffic, especially given the fact that the sum of all forms of video (TV, video on demand, Internet, and P2P) is expected to account for over 91% of global consumer traffic by 2014, whereby Internet video alone will account for 57% and 3D/HD Internet video will comprise 46% of all consumer Internet video traffic by 2014, respectively [5].

A key requirement for providing video services over FiWi access networks is to deliver the video frames in a timely manner so that the receiver can continuously play back the video. This timely video frame delivery is made challenging by the highly varying (bursty) video traffic bit rates produced by the efficient video coding standards, especially the H.264 Scalable Video Coding (SVC) standard.

2. Video MAC Protocol (VMP)

In the recent paper [6], we have taken first steps toward the development and evaluation of a suite of advanced MAC protocols, for future EPON-based R&F FiWi networks delivering streaming video. In [6] we focus on a video MAC protocol (VMP) for streaming pre-recorded video in the downstream direction. The goal of VMP is to seamlessly integrate MAC mechanisms in the optical and wireless network segments for improved performance. In our VMP protocol, we introduce three main MAC enhancement techniques to improve the quality of received video streams at the end-users, namely: (i) We examine MAC frame fragmentation in conjunction with two-level (hierarchical) frame aggregation. (ii) We introduce hybrid wireless channel access control consisting of reservation-based periods, non-polling contention-based periods, and polling contention-free periods. We achieve polling contention-free channel access through multi-polling medium access control over the integrated FiWi network segments. (iii) We introduce prefetching of video frames in conjunction with a hybrid reservation/contention-based MAC protocol over the integrated fiber-wireless network segments, whereas existing state-of-the-art MAC mechanisms consider video delivery over an isolated wireless network segment with hybrid reservation/contention-based MAC without prefetching [7].

3. Overview of Results

We found that prefetching mechanisms [8] benefit significantly from hybrid reservation/contention-based MAC protocols in FiWi networks. In particular, without reservations all video packets have to content for wireless channel access leading to a high proportion of collisions and relatively inefficient packet transport. At the other extreme, when dividing up the bottleneck bandwidth equally among the nodes receiving video streams there is no bandwidth left for contention. That is, there is effectively one "circuit" of fixed bandwidth to each receiving node and no more (global) statistical bandwidth sharing. This lack of global bandwidth sharing leads to relatively high starvation probabilities for the streaming of bursty video traffic.

The lowest playback starvation probabilities are achieved for moderate levels of reservation such that a basic level of the variable bit rate video traffic benefits from the efficient contention-free fixed-bandwidth "circuit" to each receiver. The traffic exceeding the reserved bandwidth for a node contends for the globally shared remaining bandwidth pool. Overall, our results indicate that video prefetching with channel probing reduces the video starvation probability by over an order of magnitude compared to reservation/contention MAC.

4. Summary and Outlook

We have briefly reviewed the emerging area of video streaming over Fiber-Wireless (FiWi) access networks as well as a first study of efficient integrated medium access control (MAC) mechanisms for FiWi networks with streaming prerecorded video traffic in the downstream direction. We have found that hybrid reservation/contention-based medium access control benefits significantly from prefetching of video frames with channel probing. For a wide range of reserved bandwidth levels, prefetching with channel probing robustly reduces playback starvation probabilities by over an order of magnitude.

An important direction for future work is to examine the internetworking of the FiWi network MAC protocols with metropolitan area networks, such as ring and star-based optical metro networks. Another important direction for future work is to examine efficient mechanisms for upstream (from the individual wireless stations to the Optical Line Terminal) transport of streaming video.

References

[1] N. Ghazisaidi and M. Maier, "Fiber-Wireless (FiWi) Networks: Challenges and Opportunities," IEEE Networks, vol. 25, no. 1, Jan./Feb. 2011.

[2] A. Chowdhury, H.-C. Chien, Y.-T. Hsueh, and G.-K. Chang, "Advanced System Technologies and Field Demonstration for In-Building Optical-Wireless Network With Integrated Broadband Services," IEEE/OSA Journal of Lightwave Technology, vol. 27, no. 12, pp. 1920-1927, June 2009.

[3] M. Maier, N. Ghazisaidi, and M. Reisslein, "The Audacity of Fiber-Wireless (FiWi) Networks (Invited Paper)," Proc., ICST Int. Conf. on Access Networks (AccessNets), Las Vegas, NV, Oct. 2008.

[4] P. Chowdhury, B. Mukherjee, S. Sarkar, G. Kramer, and S. Dixit, "Hybrid Wireless-Optical Broadband Access Network (WOBAN): Prototype Development and Research Challenges," IEEE Network, vol. 23, no. 3, pp. 41-48, May/June 2009.

[5] Cisco Systems, "Cisco Visual Networking Index: Forecast and Methodology, 2009-2014," June 2010.

[6] N. Ghazisaidi, M. Maier, and M. Reisslein, "VMP: A MAC Protocol for EPON-Based Video Dominated FiWi Access Networks," submitted for publication, preprint: http://mre.faculty.asu.edu/VMP.pdf

[7] R. Zhang, R. Ruby, J. Pan, L. Cai, and X. Shen, "A hybrid reservation/contention-based MAC for video streaming over wireless networks," IEEE J. Sel. Areas in Commun., vol. 28, no. 3, pp. 389–398, Apr. 2010.

[8] S. Oh, B. Kulapala, A. W. Richa, and M. Reisslein, "Continuous-time collaborative prefetching of continuous media," IEEE Transactions on Broadcasting, vol. 54, no. 1, pp. 36–52, Mar. 2008.

Navid Ghazisaidi is with Ericsson Inc., San Jose, CA. He received his MSc degree in Electrical Engineering with specialization in Telecommunications (with distinction) from the Blekinge Institute of Technology (BTH), Karlskrona, Sweden in 2006, and his PhD degree in Telecommunications from the University of Quebec, INRS, Montréal, Canada, in 2010. Dr. Ghazisaidi was a Research Associate at the School of Electrical, Computer, and Energy Engineering, Arizona State University (ASU), Tempe, AZ, USA, from September 2010 through November 2010. He was a Visiting Researcher at the

Deutsche Telekom Laboratories (T-Labs), Berlin, Germany, and participated in the prestigious European research project ACCORDANCE (A Converged Copper-Optical-Radio OFDMA-based Access Network with high Capacity and flExibility), January 2010 through April 2010. He was a Researcher in the Computer Network (CN) group at the University of Basel, Basel, Switzerland, and participated in the prestigious European research project BIONETS (BIOlogically-inspired autonomic NETworks and Services), September 2006 through January 2007. His research interests are in the area of MAC protocol design for Fiber-Wireless (FiWi) networks.

Martin Maier (maier@ieee.org) is an associate professor at the Institut National de la Recherche Scientifique (INRS), Montreal, Canada. He was educated at the Technical University of Berlin, Germany, and received MSc and PhD degrees (both with distinctions) in 1998 and 2003, respectively. In the summer of 2003, he was a postdoc fellow at the Massachusetts Institute of Technology (MIT), Cambridge. He was a visiting professor at Stanford University, Stanford, October 2006 through March 2007. Dr. Maier is a co-recipient of the 2009 IEEE Communications Society Best Tutorial Paper Award. His research activities aim at rethinking the role of optical networks and exploring novel applications of optical networking concepts and technologies across multidisciplinary domains, with a particular focus on communications, energy, and transport for emerging smart grid applications and bimodal fiber-wireless (FiWi) networks for broadband access. He is the founder and creative director of the Optical Zeitgeist Laboratory (www.zeitgeistlab.ca) and author of the book Optical Switching Networks (Cambridge University Press, 2008), which was translated into Japanese in 2009.

Martin Reisslein is an Associate Professor in the School of Electrical, Computer, and Energy Engineering at Arizona State University (ASU), Tempe. He received the Dipl.-Ing. (FH) degree from the Fachhochschule Dieburg, Germany, in 1994, and the M.S.E. degree from the University of Pennsylvania, Philadelphia, in 1996. Both in electrical engineering. He received his Ph.D. in systems engineering from the University of Pennsylvania in 1998. During the academic year 1994—1995 he visited the University of Pennsylvania as a Fulbright scholar. From July 1998 through October 2000 he was a scientist with the German National Research Center for Information Technology (GMD FOKUS), Berlin and lecturer at the Technical University Berlin. He currently serves as Associate Editor for the IEEE/ACM Transactions on Networking and for Computer Networks. He maintains an extensive library of video traces for network performance evaluation, including frame size traces of MPEG-4 and H.264 encoded video, at http://trace.eas.asu.edu. His research interests are in the areas of multimedia networking, optical access networks, and engineering education.

Prediction of Transmission Distortion for Wireless Video Communication

Zhifeng Chen and Dapeng Oliver Wu, University of Florida, FL, USA
wu@ece.ufl.edu, URL: www.wu.ece.ufl.edu

Under the prevalence of 3G/4G network and smart phones nowadays, real-time mobile video applications, e.g., videophone calls, are becoming more and more popular. However, different from the traditional video coding system, transmitting video over wireless with good quality or low end-to-end distortion is particularly challenging since the received video is subject to not only quantization error but also transmission error.

The capability of predicting transmission distortion at the transmitter can assist in designing video encoding and transmission schemes that achieve maximum video quality under resource constraints. Specifically, transmission distortion prediction can be used in the following three applications in video encoding and transmission: 1) mode decision, which is to find the best intra/inter-prediction mode for encoding a macroblock (MB) with the minimum rate-distortion (R-D) cost given the instantaneous PEP, 2) cross-layer encoding rate control, which is to control the instantaneously encoded bit rate for a real-time encoder to minimize the frame-level end-to-end distortion given the instantaneous PEP, e.g., in video conferencing, 3) packet scheduling, which chooses a subset of packets of the pre-coded video to transmit and intentionally discards the remaining packets to minimize the GOP-level (Group of Picture) end-to-end distortion given the average PEP and average burst length, e.g., in streaming pre-coded video over networks. All the three applications require a formula for predicting how transmission distortion is affected by their respective control policy, in order to choose the optimal mode or encoding rate or transmission schedule.

However, predicting transmission distortion poses a great challenge due to the spatio-temporal correlation inside the input video sequence, the nonlinearity of both the encoder and the decoder, and varying PEP in time-varying channels. The existing transmission distortion models can be categorized into the following three classes: 1) pixel-level or block-level models (applied to prediction mode selection) [1,2,3]; 2) frame-level or packet-level or slice-level models (applied to cross-layer encoding rate control) [4,5,6,7,8]; 3) GOP-level or sequence-level models (applied to packet scheduling) [9,10,11,12,13]. Although different transmission distortion models work at different levels, they share some common properties, which come from the inherent characteristics of wireless video communication system, that is, spatio-temporal correlation, nonlinear codec and time-varying channel. However, none of these existing works analyzed the effect of non-linear clipping noise on the transmission distortion, and therefore cannot provide accurate transmission distortion estimation.

In our work [14], transmission distortion is analytically derived, for the first time, as a closed-form function of packet error probability (PEP), video frame statistics, and system parameters. To support real-time mobile video applications, In our work [16], practical algorithms to estimated frame-level transmission distortion (FTD) and pixel-level transmission distortion (PTD) are derived based on the formulae derived in [14]. The methodology is applicable to deriving formulae for slice/packet/GOP-level distortion.

In estimating frame-level transmission distortion, the system parameters, e.g., correlation ratio and propagation factor, in the FTD model are estimated from instantaneous video frame statistics and channel conditions. This facilitates the design of a low complexity algorithm, called RMPC-FTD algorithm [16], for estimation FTD. In addition, since FTD can be estimated from instantaneous video frame statistics and channel conditions, the FTD estimation algorithm are applicable to time-varying channel, that is, the error processes could be non-stationary. However, existing FTD algorithms estimate their parameters by using the statistics averaged over multiple frames and assume these statistics do not change over time; their models all assume the error process is stationary. In contrast, due to the use of instantaneous video frame statistics and channel conditions, the RMPC-FTD algorithm proposed in [16] is more accurate than existing FTD algorithms for real-time video communication.

The estimation algorithm for PTD is similar to the FTD estimation algorithm since the PTD formula is a special case of the FTD formula as discussed in [14]. However, in some existing video encoders, e.g., H.264 reference code [15], motion estimation and prediction mode decision are separately considered. Therefore, the MV and corresponding residual are known for distortion estimation in mode decision. In such a case, the PTD estimation algorithm can be simplified with known values of the MV and corresponding residual, compared to using their statistics. In [16], a PTD estimation algorithm, called RMPC-PTD, is developed for such a case.

In existing video encoders, prediction mode decision is to choose the best prediction mode in the sense of minimizing the Rate-Distortion (R-D) cost for each MB or sub-MB. Estimation of the pixel level end-to-end distortion (PEED) for different prediction modes is needed for calculating R-D cost in prediction mode decision. In [16], RMPC-PTD is further extended to support estimating PEED; the resulting algorithm is called RMPC-PEED. Compared to ROPE algorithm [1], RMPC-PEED algorithm [16] is more accurate since the significant effect of clipping noise on transmission distortion is considered. Another advantage over ROPE algorithm is that RMPC-PEED algorithm is much easier to be extended to support averaging operations, e.g., interpolation filter. Compared to LLN algorithm [2,3], the computational complexity and memory requirement of RMPC-PEED algorithm are much lower and the estimated distortion has smaller variance.

Experimental results [16] show that applying the RMPC-PEED algorithm in prediction mode selection in H.264 encoder [15] achieves an average PSNR gain of 1.44dB over ROPE algorithm for "foreman" sequence under PEP=5%; and it achieves an average PSNR gain of 0.89dB over LLN algorithm for "foreman" sequence under PEP=1%.

References

[1] R. Zhang, S. L. Regunathan, and K. Rose, "Video coding with optimal inter/intra-mode switching for packet loss resilience," IEEE Journal on Selected Areas in Communications, vol. 18, no. 6, pp. 966–976, Jun. 2000.

[2] T. Stockhammer, M. Hannuksela, and T. Wiegand, "H. 264/AVC in wireless environments," IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 657–673, 2003.

[3] T. Stockhammer, T. Wiegand, and S. Wenger, "Optimized transmission of h.26l/jvt coded video over packet-lossy networks," in IEEE ICIP, 2002.

[4] M. Sabir, R. Heath, and A. Bovik, "Joint source-channel distortion modeling for MPEG-4 video," IEEE Transactions on Image Processing, vol. 18, no. 1, pp. 90–105, 2009.

[5] K. Stuhlmuller, N. Farber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," IEEE Journal on Selected Areas in Communications, vol. 18, pp. 1012–1032, Jun. 2000.

[6] J. U. Dani, Z. He, and H. Xiong, "Transmission distortion modeling for wireless video communication," in Proceedings of IEEE Global Telecommunications Conference (GLOBECOM'05), 2005.

[7] Z. He, J. Cai, and C. W. Chen, "Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding," IEEE Transactions on Circuits and System for Video Technology, special issue on wireless video, vol. 12, pp. 511–523, Jun. 2002.

[8] Y. Wang, Z. Wu, and J. M. Boyce, "Modeling of transmission-loss-induced distortion in decoded video," IEEE Transactions on Circuits and Systems for Video Technology, vol. 16, no. 6, pp. 716–732, Jun. 2006.

[9] J. Chakareski, J. Apostolopoulos, W.-T. Tan, S. Wee, and B. Girod, "Distortion chains for predicting the video distortion for general packet loss patterns," in Proc. ICASSP, 2004.

[10] J. Chakareski, J. Apostolopoulos, S. Wee, W.-T. Tan, and B. Girod, "Rate-distortion hint tracks for adaptive video streaming," IEEE Transactions on Circuits and Systems for Video Technology, vol. 15, no. 10, pp. 1257–1269, 2005.

[11] C. Zhang, H. Yang, S. Yu, and X. Yang, "GOP-level transmission distortion modeling for mobile streaming video," Signal Processing: Image Communication, 2007.

[12] M. T. Ivrlaˇ c, L. U. Choi, E. Steinbach, and J. A. Nossek, "Models and analysis of streaming video transmission over wireless fading channels," Signal Processing: Image Communication, vol. 24, no. 8, pp. 651–665, Sep. 2009.

[13] Y. J. Liang, J. G. Apostolopoulos, and B. Girod, "Analysis of packet loss for compressed video: Effect of burst losses and correlation between error frames," IEEE Transactions on Circuits and Systems for Video Technology, vol. 18, no. 7, pp. 861–874, Jul. 2008.

[14] Z. Chen and D. Wu, "Prediction of Transmission Distortion for Wireless Video Communication: Part I: Analysis," 2010, http://www.wu.ece.ufl.edu/mypapers/journal-1.pdf.

[15] "H.264/AVC reference software JM16.0," Jul. 2009. [Online]. Available: http://iphome.hhi.de/suehring/tml/download

[16] Z. Chen, D. Wu, "Prediction of Transmission Distortion for Wireless Video Communication: Algorithm and Application," Journal of Visual Communication and Image Representation, vol. 21, no. 8, pp. 948-964, November 2010.

Zhifeng Chen received the B.E. degree from the East China University of Science and Technology, Shanghai, China, in 2001, and the M.S. and Ph.D. degrees from the University of Florida, Gainesville, Florida, in 2008 and 2010 respectively. From 2002 to 2003, he was an engineer in EPSON (China), and from 2003 to 2006, he was a senior engineer in Philips (China), both working in mobile phone system solution design. From May 2009 to Aug 2009, he was an intern in Dolby, Burbank, CA, where he had worked in error resilient rate distortion optimization. He joined the strategic engineering department at Interdigital, King of Prussia, PA, in 2010, where he is currently a staff engineer working in the video coding research. His research interests include low-complexity video and image compression, perceptual video coding, Error-resilient video coding, rate-distortion optimization, rate control, cross-layer design, information theory, statistics, signal processing. He is the author of several journal and conference papers and has been awarded four patents.

Dapeng Oliver Wu received Ph.D. in Electrical and Computer Engineering from Carnegie Mellon University, Pittsburgh, PA, in 2003.

Since 2003, he has been on the faculty of Electrical and Computer Engineering Department at University of Florida, Gainesville, FL, where he is currently Associate Professor. His research interests are in the areas of networking, communications, signal processing, computer vision, and machine learning. He received University of Florida Research Foundation Professorship Award in 2009, AFOSR Young Investigator Program (YIP) Award in 2009, ONR Young Investigator Program (YIP) Award in 2008, NSF CAREER award in 2007, the IEEE Circuits and Systems for Video Technology (CSVT) Transactions Best Paper Award for Year 2001, and the Best Paper Award in International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks (QShine) 2006.

Currently, he serves as an Associate Editor for IEEE Transactions on Wireless Communications, IEEE Transactions on Circuits and Systems for Video Technology, Journal of Visual Communication and Image Representation, and International Journal of Ad Hoc and Ubiquitous Computing. He was the founding Editor-in-Chief of Journal of Advances in Multimedia between 2006 and 2008, and an Associate Editor for IEEE Transactions on Vehicular Technology between 2004 and 2007. He is also a guest-editor for IEEE Journal on Selected Areas in Communications (JSAC), Special Issue on Cross-layer Optimized Wireless Multimedia Communications. He will serve as Technical Program Committee (TPC) Chair for IEEE INFOCOM 2012, and has served as TPC Chair for IEEE International Conference on Communications (ICC 2008), Signal Processing for Communications Symposium, and as a member of executive committee and/or technical program committee of over 50 conferences. He has served as Chair for the Award Committee, and Chair of Mobile and wireless multimedia Interest Group (MobIG), Technical Committee on Multimedia Communications, IEEE Communications Society.

MSE cross-layer optimization criteria: What else?

Laura Toni, Italian Institute of Technology, Genova, Italy
Pascal Frossard, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
laura.toni@iit.it, pascal.frossard@epfl.ch

1. Introduction

Recently, video streaming over wireless networks has become quite popular with the development of novel exciting mobile media applications. It brings interesting research challenges with a deep change in design paradigms and research emphases compared to traditional streaming applications. In particular, cross-layer optimization techniques have been under intense research as primary strategies for adaptation to dynamic channel conditions, since they permit to improve the quality of experience by optimizing the network architectures across traditional OSI stack layers.

The main issue to be faced when adopting a cross-layer approach is the definition of a reliable metric to formulate the quality optimization problem. Ideally, a good metric should be straightforward to understand and immediate to calculate, encompassing, at the same time, the different components that influence end-user video distortion in wireless streaming. At the source coder, the distortion typically decreases with increasing rate, but the actual rate-distortion characteristic depends on the video content. For example, in a variable rate multiuser scenario, an accurate quality assessment can help the video senders in optimizing their transmission settings (packet dropping prior the transmission or packet scheduling) that, in turn, results in efficient usage of the network resources towards minimal average distortion. Additionally, a good distortion metric should also be able to capture the effects of channel impairments on the received video, or equivalently the distortion degradation due to packet losses and error concealment. Typically, an increase in the error protection tends to reduce the impact of losses, but this distortion characteristic is also influenced by the video content. In general, one faces an optimization problem where bits have to be traded off between source coding rate and channel protection for optimized video quality. An accurate measure of video distortion becomes key in the selection of the appropriate optimization techniques.

The video quality assessment might be considered from two different angles: as a mathematical evaluation (*mathematical metric*) or as a visual-perception test (*visual perception metric*). While the former is simple to assess and easy to reproduce, the latter provides results that are often more complex to evaluate but also more closely related to the video quality experienced by the final user, *i.e.*, the human eye. The compelling question is “Which is the best distortion metric to be considered for cross-layer design?” [2, 3]. In this letter, we aim at providing an insight on this topic, highlighting the advantages and limits of both mathematical and visual perception metrics.

2. Mathematical Metrics

Objective tests evaluate the video quality through criteria that can be rigorously measured and automatically computed. Among these mathematical metrics, the most common is the mean square error (MSE) metric. The MSE criterion has many attractive features. First, it is straightforward to evaluate. Being x_i and y_i the i-th pixel of the original and reconstructed frames of the video sequence, respectively, the MSE is calculated as the mean energy of the error signal $e=(x_i-y_i)$. In addition, MSE is able to capture the effects of both source and channel distortions in the video degradation. Finally, minimum-MSE optimization problem usually leads to closed-form analytical solutions or iterative numerical procedures that are easy to implement. For all these reasons, MSE is widely used as a parameter to be minimized in cross-layer design. [5-8].

For example, the authors propose a cross-layer algorithm for distributed video rate allocation over wireless systems in [5, 6]. The MSE distortion metric is justified by previous studies on subjective test [9], which have demonstrated that the rate allocation algorithm aimed at minimizing the overall MSE agrees sufficiently well with subjective opinions. MSE distortion is minimized also in [8], where a Rate-Distortion (RD) based video packet selection for distributed streaming over a shared communication channel is proposed. Although optimizing MSE does not necessarily correspond to optimizing subjective quality, the authors demonstrate that a video-aware cross-layer technique achieves a lower received video distortion in terms of MSE, compared to baseline algorithms.

A comprehensive RD analysis of video quality due to source and channel distortion has been reported in [10-12]. MSE typically decreases non linearly with the increase of the received rate. It can be modeled with parametric functions, where the parameters of the model can be evaluated through non linear regression techniques or least-square methods. Even if complete MSE models might be constructed, the proper handling of the temporal evolution of the video quality remains unsolved. Two different video streams with the same average distortion might not be perceived equally from an end-user point of view. This is mainly caused by the fact that average MSE distortion cannot distinguish the frame-by-frame quality variation and its effect on the final user. These effects are difficult to capture in cross-layer streaming solutions that usually try to minimize the average distortion.

Another important aspect that should improve the fidelity of the distortion metric is the relation between MSE and packet losses. In [13], the authors analyze the evolution of expected distortion versus loss pattern, for video communication over error-prone channels. They verify that the packet loss pattern, and in particular the burst length, does have a significant effect on the resulting distortion; hence it should be taken into consideration in the optimization problems. Despite its importance in video cross-layer designs, the relation between distortion and channel impairment usually depends on many parameters, such as coding scheme, the bit rates and the network architectures. Even if MSE is a priori a simple metric, its appropriate use in the design of cross-layer optimization solutions is not as trivial as it appears.

3. Visual Perception Metrics

Since the ultimate receivers of wireless video streams are usually human observers, mathematical metrics, which measure the video distortion without considering perceptual factors, might not reflect the video quality perceived by the human eye. Visual perception metrics (e.g., visual quality metric, VQM, [14], or subjective tests) are able to capture the effects of source distortion as well as channel distortion (loss distortion, temporal error propagation, variable bit-rates) with high fidelity to the human perception, at a price of complexity in the evaluation. Much of the efforts in understanding the visual impact of channel impairments are focused on the evaluation of the overall perceived quality. In [14], a subjective study to assess the visual quality of compressed video affected by wireless channel impairments is provided. The main limit is that the evaluation of mean opinion score (MOS) is extremely complex (e.g., the quality scales are generally not equally interpreted by the users). It further requires fully decoded video streams and is highly dependent on the testing conditions. It is not easy to factor the effects of each system parameters in terms of overall subjective quality.

To the best of our knowledge, there exists no objective metric which is able to assess the visual quality of wireless video in terms of the relative importance of the different packets to be employed in cross-layer optimization techniques. An alternative consists in taking into account a set of simpler metrics, which can help the system designer in optimizing the transmission scheme in order to augment the visual quality. One of these metrics is the bitstream-based metric provided in [15, 16]. Bitstream-based metrics predict video quality looking at the packet header information and the encoded bitstream without decoding the video source. In these works, the metric aims at evaluating the effects of packet losses on the perceived quality. The authors conducted subjective tests in which the viewers' task is to indicate when an artifact, due to packet losses, becomes visible. The subjective tests are conducted for video streams in which isolated losses could be experienced. From these tests, a packet loss visibility metric (PLV) has been proposed with the goal of predicting whether a packet loss in the video stream becomes visible from a human viewer. The main advantages of the proposed technique are that the PLV is predictable with an objective metric, takes into account the temporal evolution of the video quality, and does not require to fully decode the video sequence. The authors demonstrate that the adoption of the PLV model as packet prioritization policy for video transmission leads to a better overall perceived quality compared to baseline cross-layer streaming solution.

The main drawback of the PLV model lies on a simplistic channel model with isolated losses. Extensions of visibility metrics to more realistic wireless channel models should however lead to improved cross-layer design as long as the computational complexity stays limited. Fidelity and complexity are therefore posed as a trade-off in the selection of the distortion metric in cross-layer design optimization.

4. Conclusions

In this letter, we provide an insight on the distortion metric that could be adopted in order to optimize the cross-layer video system design. Mathematical metrics, especially the MSE, are simple to evaluate and well accepted from the video processing community. Although reducing the MSE is not equivalent to minimizing the visual quality, the MSE metric provides a reasonable level of fidelity when dealing with cross-layer techniques, but only provides a coarse approximation of the actual visual distortion. Visual quality metrics, which estimate the actual quality perceived by the final user, are suitable to benchmark systems and algorithms in comparative studies, but their use in cross-layer design is not trivial. Recent works on visibility models offer promising perspectives in cross-layer design, as long as these models are able to adapt to realistic wireless channels with a controllable complexity.

References

[1] S. Mohamed and G. Rubino, "A study of real-time packet video quality using random neural networks," IEEE Transactions on Circuits and Systems for Video Technology, vol. 12, pp. 1071-1083, 2002.

[2] W. Zhou and A. C. Bovik, "Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures," IEEE Signal Processing Magazine, vol. 26, pp. 98-117, 2009.

[3] S. Winkler and P. Mohandas, "The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics," IEEE Transactions on Broadcasting, vol. 54, pp. 660-668, 2008.

[4] R. R. Pastrana-Vidal and J.-C. Gicquel, "Automatic quality assessment of video fluidity impairments using a no-reference metric" Int'l Workshop on Video Proc. and Quality Metrics, Jan. 2006.

[5] Z. Xiaoqing and B. Girod, "Distributed Media-Aware Rate Allocation for Wireless Video Streaming," IEEE Transactions on Circuits and Systems for Video Technology, vol. 20, pp. 1462-1474.

[6] Z. Liang, W. Xinbing, T. Wei, G. M. Muntean, and B. Geller, "Distributed scheduling scheme for video streaming over multi-channel multi-radio multi-hop wireless networks," IEEE Journal on Selected Areas in Communications, vol. 28, pp. 409-419.

[7] Y.S. Chan, P. C. Cosman, and L. B. Milstein, "A Multiple Description Coding and Delivery Scheme for Motion-Compensated Fine Granularity Scalable Video," IEEE Transactions on Image Proc., vol. 17, pp. 1353-1367, 2008.

[8] J. Chakareski and P. Frossard, "Rate-distortion optimized distributed packet scheduling of multiple video streams over shared communication resources," IEEE Transactions on Multimedia, vol. 8, pp. 207-218, 2006.

[9] Z. Xiaoqing and B. Girod, "Subjective evaluation of multi-user rate allocation for streaming heterogeneous video contents over wireless networks," in IEEE ICIP 2008. 15th 2008, pp. 3092-3095.

[10] P. Pahalawatta, R. Berry, T. Pappas, and A. Katsaggelos, "Content-Aware Resource Allocation and Packet Scheduling for Video Transmission over Wireless Networks," IEEE Journal on Selected Areas in Communications, vol. 25, pp. 749-759, 2007.

[11] C. Yim and A. C. Bovik, "Evaluation of temporal variation of video quality in packet loss networks," Signal Processing: Image Communication, vol. 26, pp. 24-38.

[12] M. Naccari, M. Tagliasacchi, and S. Tubaro, "No-Reference Video Quality Monitoring for H.264/AVC Coded Video," IEEE Transactions on Multimedia, vol. 11, pp. 932-946, 2009.

[13] Y. J. Liang, J. G. Apostolopoulos, and B. Girod, "Analysis of Packet Loss for Compressed Video: Effect of Burst Losses and Correlation Between Error Frames," IEEE Transactions on Circuits and Systems for Video Technology, vol. 18, pp. 861-874, 2008.

[14] A. K. Moorthy, K. Seshadrinathan, R. Soundararajan, and A. C. Bovik, "Wireless Video Quality Assessment: A Study of Subjective Scores and Objective Algorithms," IEEE Transactions on Circuits and Systems for Video Technology, vol. 20, pp. 587-599.

[15] S. Kanumuri, P. C. Cosman, A. R. Reibman, and V. A. Vaishampayan, "Modeling packet-loss visibility in MPEG-2 video," IEEE Transactions on Multimedia, vol. 8, pp. 341-355, 2006.

[16] L. Ting-Lan, S. Kanumuri, Z. Yuan, D. Poole, P. C. Cosman, and A. R. Reibman, "A Versatile Model for Packet Loss Visibility and its Application to Packet Prioritization," *IEEE Transactions on Image Processing,* vol. 19, pp. 722-735.

Laura Toni has a postdoctoral position at the TeleRobotics and Applications (TERA) department at the Italian Institute of Technology (IIT), Genova, Italy. In 2009, she received her Ph.D. degree in electronics, computer science and telecommunications from the University of Bologna, Italy. During 2007, she had been a visiting scholar at the University of California at San Diego, CA, working on video processing over wireless systems. Since 2009, she has been a frequent visitor to the UCSD, working on joint source and channel coding on wireless communication systems. Her research interests are in the areas of image and video processing, wireless communications, and underwater communications.

Pascal Frossard received the M.S. and Ph.D. degrees, both in electrical engineering, from the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, in 1997 and 2000, respectively. Between 2001 and 2003, he was a member of the research staff at the IBM T. J. Watson Research Center, Yorktown Heights, NY. Since 2003, he has been an assistant professor at EPFL, where he heads the Signal Processing Laboratory (LTS4). His research interests include image representation and coding, visual information analysis, distributed image processing and communications, and media streaming systems.

Video-on-Demand Broadcasting to Mobile Wireless Users

Duc A. Tran,University of Massachusetts, Boston, MA
duc.tran@umb.edu

1. Introduction

Today's wireless technologies such as IEEE 802.16 (a.k.a., WiMAX) for long-haul communications and IEEE 802.11 (e.g., WiFi) and Bluetooth for short distances are widely deployed. As beneficiaries, users can move freely without disconnection from the network and can enjoy ubiquitous entertainment services. This paper briefly discusses technologies that enable mobile video-on-demand (VOD) services. Unlike other video services such as pay-per-view (PPV) and video in demand (VID), individual VOD clients in an area are able to watch different programs whenever they wish to, not just in time as in VID or pre-scheduled as in PPV.

2. VoD on the Internet

VOD services are already available on the Internet. News video clips can be rendered on demand on most Web media outlets (e.g., cnn.com, espn.com). Video commercials in business areas ranging from automotive to real estate to health and travel can also be played on demand via a product of Comcast called ComcastSpotlight [1]. The major server providers for VOD deployment include Motorola On-Demand Solutions [2], SeaChange International [3], and Concurrent Corp. [4]. Informa Telecom (http://www.informatm.com) predicted a revenue of more than 10.7 billion US dollars from VOD services offered to more than 350 million households by 2010.

The designs for current VOD systems can be categorized into three main approaches: client/server, peer-to-peer, and periodic broadcast. These approaches are like “apple and orange” when it comes to comparison because each approach is effective for a certain subset of VOD applications. Video popularity is known to practically follow an 80/20-like rule of thumb; that is, most clients would be interested in only a few popular videos. As such, periodic broadcast should be best used for transmitting popular videos to a large number of clients, while client/server and P2P techniques are better suitable for non-popular videos or for videos requested by not too many clients.

3. VoD for Mobile Wireless Users

When deployed to a wireless environment, the success of existing VOD designs may not apply. Because the most popular form of wireless communications in a local area is by IEEE 802.11 technologies, the network bandwidth shared by all the users covered by an access point is typically limited, putting a cap on the number of video streams that can be delivered simultaneously. Taking into account signal inference, the effective number of such streams would actually be much fewer. Also, an 802.11-enabled host can only reach other devices within 100m of its radius, while that radius is 10m for Bluetooth; if a user is too far away from any access point, how can it get the video service?

Fortunately, today’s wireless hosts are able to concurrently participate in multiple connections: with the access point in the infrastructure-based mode and with a nearby host in the ad hoc mode. Therefore, it is possible that a distant user could get the video service from the access point via several intermediate hosts. The problem is that significant amounts of bandwidth and energy of the intermediate mobile hosts are consumed. We do not have this problem with the typical Internet.

So arise two questions: 1) what should be the architecture for a mobile wireless VOD system? and 2) what should be the communication protocol for a client to download a video from the video server?

To cope with the limited wireless bandwidth, as each wireless transmission is a broadcast where every host can hear, it is natural to think that it would be more efficient if we adopt the broadcast approach. It would be best if we apply the broadcast approach to the most popular videos and the client/server approach only for ad hoc unpopular video requests.

To cope with the limited wireless coverage, we should allow sharing of video contents among the users. For instance, instead of playing a video through multiple hops from an access point, we hope to play the video or part of it in some existing users nearby. In other words, we should adopt the P2P approach for this kind of users.

4. MobiVOD

Despite many periodic broadcast designs for the Internet, they may not be directly applicable to a wireless network due to requirements on client bandwidth and caching space (see Table 1).

Table 1. Client Resource Requirement for Representative Periodic Broadcast (PB) Techniques

Technique	Caching Space (% video size)	Bandwidth (times playback rate)
Staggered [5]	0	1
Skyscraper [6]	10	2
Pyramid [7]	75	>=4
Permutation [8]	20	>=2
Pagoda [9]	45	>=5
Harmonic [10]	40	>=5
Fast [11]	50	>=6

None of these periodic broadcast techniques can provide true VOD because their service delay is the duration of the first segment. MobiVOD [12] is a mobile wireless near-true-VOD solution based on Staggered Broadcasting (SB). SB is chosen because of its modest resource requirement on the client side, thus suitable for clients of mobile wireless networks. MobiVOD erases SB's service delay by leveraging video content sharing between the wireless clients in a P2P manner

The system architecture for MobiVOD consists of three components: video server, clients, and local forwarders. Because it is not possible for the server to wirelessly transmit a video to clients located in a too-wide geographic area, local forwarders are deployed to relay the video broadcast (using SB) from the server to clients in their corresponding service area. The communication between local forwarders and the server is via wired broadband or wireless like WiMAX. A local forwarder broadcasts video content to its local coverage using a short-range wireless technology, for example, IEEE 802.11.

When a new client starts the video request and the first segment is not yet available on any broadcast channel, the new client can get this segment instantly from a nearby client who has a cache of it. Obviously, if the cache is several hops away, it is helpless because there is no efficient way for the new client to download the segment in a multi-hop manner. A key property of MobiVOD is its attempt to make the cache available within one hop (i.e., from a direct neighbor), without asking too many nodes to cache. The idea for MobiVOD to do this is to require only the clients that belong to a dominating set of the clients to cache the first segment. Simulation results are provided in [12], showing a service delay nine times better than that of SB in most scenarios.

Although originally designed to work with SB, MobiVOD can be modified to work with most other PB techniques. Also, in real-world implementation, we actually have to deal with different types of clients, especially those having bandwidth less than the video consumption rate. For this purpose, one can employ a multi-resolution or layered video coding approach [13, 14]. Each video can be encoded into several "layers", including a base layer and one or more enhancement layers. The base layer provides the version of least quality, while its combination with enhancement layers provides incrementally better quality. These layers are broadcast on separate channels. A new client selects a combination of layers that best match its resource constraints and only tunes in the corresponding channels to download such layers. As for the initial segment that the client misses from the current broadcasts, it searches for a nearby client who caches a "version" of the first segment (a version is a combination of the base layer and one or more enhancement layers). If more than one such client are found, the client with the highest-quality version is selected.

5. Conclusions

While broadcasting is the nature of wireless communications, VOD broadcasting is not trivial to be implemented in a wireless network. This is because, unlike the Internet where one-to-one communication does not affect nodes that do not involve in the communication, wireless transmission between two nodes may interfere with transmissions by other nodes. MobiVoD is a feasible VOD technique for wireless environments, offering much better VOD feel when combined with a periodic broadcasting technique. Using today's technologies, clients can communicate wirelessly through access points, base stations, or one-to-one in an ad hoc manner with each other. The key idea of MobiVoD is to leverage client collaboration.

The future of VOD is bright. A VOD system of the future will be realization of the video rental shop brought into the home, and wherever the client goes. Airlines could provide VOD services in airport lounges to entertain passengers on their own PDA while they are waiting for a flight; a museum could provide video information on the exhibits on demand over the wireless network; in education, a university could also install such a system on campus to allow students to watch video recorded earlier from lectures they were not able to attend. Despite all these potential demands, research and development on mobile wireless VOD remain sparse and it will be interesting (and rewarding) to investigate further into this problem.

References

[1] Comcast spotlight: Video advertising on demand. http://www.comcastspotlight.com.
[2] Motorola on-demand solutions. http://broadband.motorola.com/business/ondemand/VideoOnDemand.html.
[3] Seachange international: The global leader in video on demand. http://www.schange.com/.
[4] Concurrent. On-demand and real-time linux solutions. http://www.ccur.com/.
[5] J. B. Kwon and H. Y. Heom. Providing vcr functionality in staggered video broadcasting. IEEE Transactions on Consumer Electronics}, 48(1):41--48, 2002.
[6] K. A. Hua and S. Sheu. Skyscraper broadcasting: A new broadcasting scheme for metropolitan video-on-demand systems. In Proc. of the ACM SIGCOMM'97, pages 89--100, Cannes, France, Sepetember 1997.
[7] S. Viswanathan and T. Imielinski. Metropolitan area video-on-demand service using pyramid broadcasting. ACM Multimedia systems Journal}, 4(4):179--208, August 1996.
[8] C. C. Aggarwal, J. L. Wolf, and P. S. Yu. A permutation-based pyramid broadcasting scheme for video-on-demand systems. In Proc. of the IEEE Int'l Conf. on Multimedia Systems '96, pages 118--126, Hiroshima, Japan, June 1996.
[9] J. F. Paris, S. W. Carter, and D. D. E. Long. A hybrid broadcasting protocol for video on demand. In ACM/SPIE Conference on Multimedia Computing and Networking , 1999.
[10] L. Juhn and L. Tseng. Harmonic broadcasting for video-on-demand service. IEEE Transactions on Broadcasting, 43(3):268--271, 1997.
[11] L. Juhn and L. Tseng. Fast data broadcasting and receiving scheme for popular video service. IEEE Transactions on Broadcasting}, 44(1):100--105, 1998.
[12] D. A. Tran, K. A. Hua, and M. Le. MobiVOD: A video-on-demand system design for mobile ad hoc networks. In Proceedings of IEEE International Conference on Mobile Data Management (MDM 2004), Berkeley, CA, USA, January 19-22 2004.
[13] S. Bajaj, L. Breslau, and S. Shenker. Uniform versus priority dropping for layered video. In ACM SIGCOMM, pages 131--143, 1998.
[14] V. N. Padmanabhan, H. J. Wang, P. A. Chou, and K. Sripanidkulchai. Distributing streaming media content using cooperative networking. In ACM/IEEE NOSSDAV, pages 177--186, Miami, FL, USA, May 12-14, 2002.

Duc A. Tran is an Assistant Professor in the Department of Computer Science at the University of Massachusetts at Boston, where he leads the Network Information Systems Laboratory. He received a PhD degree in Computer Science from the University of Central Florida (Orlando, Florida). Dr. Tran's general interests are in the areas of networking and distributed systems, particularly in support of information systems that can scale with both network size and data size.. The results of his work have led to research grants from the National Science Foundation (NSF) and two Best Papers (ICCCN 2008, DaWak 1999). Also, his contribution on P2P streaming is widely-cited in this area (to date approaching 600 citations according to Google Scholar). Dr. Tran has engaged in many professional activities, serving as a Review Panelist for the NSF, Editorial Board member for the Journal on Parallel, Emergent, and Distributed Systems (2010-date) and ISRN Communications Journal (2010-date), Guest-Editor for the Journal on Pervasive Computing and Communications (2009), TPC Co-Chair for CCNet 2010, GridPeer (2009, 2010, 2011), and IRSN 2009, TPC Vice-Chair for AINA 2007, and TPC member for 40+ international conferences. He is a member of the ACM and the IEEE.

TECHNOLOGY ADVANCES

Interactive and Smart Multimedia Services

Guest Editor: Min Chen, Seoul National University, Korea
minchen@ieee.org

Today with the fast developments in electronics industry and the amazing growing of the potential customers, media are increasingly ubiquitous: more and more people live in a world of Internet pop-ups and streaming television, mobile phone texting and video clips, etc. These ubiquitous media become more and more closer to our life experience, enabling brand environments for interactive and smart multimedia services, which makes it possible to access various multimedia contents everywhere at any time. Thus, the design of interactive multimedia and services combines aspects of ubiquitous networking, ambient intelligence, embedded multimedia systems, and smart environments, which has drawn substantial attention from the research community.

In the first paper "Shelter: Smartphone Bridged Socialized Body Networks for Epidemic Control" by X. Bai, H. Wang and H. Fang uses body area networks, social networks and smartphone to tackle epidemic control challenges. The major innovation in this approach is to develop a smartphone bridged social body networks to identify the critical networks that represent epidemic patterns. The proposed approach fills the gap between social information collection and body including external environment information collection, providing a promising solution for real time epidemic control.

Jiangchuan Liu et al., in "Frequency-Aware Adaptive Indexing for P2P VoD Service with Frequent Seeks", address the key challenge, chunk indexing, for VoD application with frequent interactivities. The authors propose a novel indexing scheme which realizes the quick and low-cost chunk discovering. The major idea is to adapt itself to the user request pattern adaptively. The knowledge extracted from the overlay facilitates the prefetching mechanism efficiently. The preliminary simulation results fully demonstrate its superiority in supporting VoD service, especially for non-uniform user access pattern.

The paper, titled "LCHQ: Low Complexity with High Quality Video Codec for the Mobile Multimedia Communications" by X. Liu and J. Ma, addresses the current situations and bottleneck of the high-quality video coding standard for the mobile multimedia communications. The authors focus on the dominant video coding standards and analyze their serious complexity problem which will take trouble to the mobile devices, such as energy life, processing time and so on. They also discuss the development directions and improvement challenges of the future LCHQ video coding standard for the mobile communication systems. This article provides relevant summary and useful outlook for the future related research.

In the last paper "3D Motion-Sensing Interactive HMI for Applications in Multimedia Services" by Yueh-Min Huang et al., mainly addresses on interactive multimedia services which interacts singly according to respective actions from the user still belong to an one-way interactive system. However users do not feel the realism with this type of interactive service. This paper introduces gloves implemented with infrared signal LED and triple-axis accelerometer for hand motion recognition. The system receives the response when an item is attained for service. The physical reaction is presented with small motors and springs to increase the tension display within the glove. This system already been successfully implemented in normal operating systems and provided an innovative idea for interactive multimedia service.

Research in interactive and smart multimedia services has seen both challenges and opportunities with emerging new applications and systems. It is intended to be a novel efficient and effective component for the Internet of Things, and will greatly impact everyone's life in every facet, including entertainment, socialization, business, healthcare and education. We would like to thank all the authors for their contributions and hope these articles can provide some research interests in this area.

Min Chen is an assistant professor in School of Computer Science and Engineering at Seoul National University (SNU). Before joining SNU, he was a Post-Doctoral Fellow and Research Associate at University of British Columbia for three and half years. He received the Best Paper Runner-up Award fr om QShine 2008. He has published more than 120 technical papers. Dr. Chen is the author of OPNET Network Simulation (Tsinghua University Press, 2004). He serves as TPC co-chair for BodyNets-2010, workshop co-chair for Chinacom 2010 and 2011. He is the co-chair of MMASN 2009, UBSN 2010 and NCAS 2011. He is workshop co-chair for EMC 2010. He is co-chair for ASIT 2010 and 2011. He was the TPC chair of ASIT-09 and GCCN-10, TPC co-chair of PCSI-09, publicity co-chair of PICom-09. He serves as editor or associate editor for Wireless Communications and Mobile Computing (WCMC), IET Communications, Wiley I. J. of Security and Communication Networks (SCN), Journal of Internet Technology (JIT), KSII Transactions on Internet and Information Systems (TIIS), International Journal of Sensor Networks (IJSNet), etc. He is managing editor for IJAACS. Currently, his research focuses on multimedia and communications, such as multimedia transmission over wireless network, wireless sensor networks, body sensor networks, RFID, ubiquitous computing, intelligent mobile agent, pervasive computing and networks, E-healthcare, medical application, machine to machine communications and Internet of Things, etc. He is an IEEE Senior Member.

Shelter: Smartphone Bridged Socialized Body Networks for Epidemic Control
(invited paper)

Xiaole Bai, Honggang Wang, University of Massachusetts Dartmouth, MA, USA
{xbai, hwang1}@umassd.edu
Hua Fang, University of Massachusetts Medical School, Worcester, MA, USA
hua.fang@umassmed.edu

We propose using information, computing and networking innovations to tackle epidemic control challenges.

1. Motivation

The explosive growth of human mobility and social structure complexity in the past century, accompanied with population increment, has led to a variety of new problems related to public health. As a prominent challenge, the high density of people and their intensive social interaction lead to a significant risk of epidemics. This was evident in the recent outbreak of SARS, H1N1, and H5N1. Besides the high risk of natural disease, such social reality and trend also result in a significant risk of human initiated bio-terrorism events, such as a smallpox based bio-terror attacks [1].

Outbreak and epidemic control requires swift and accurate action. However, the existing approaches are far from sufficient, because they can not support ***real-time critical information collection and analysis***. They are the key to the success of accurate online decision making since evolution or mutation can happen in virus and pathogens [2] and social structure can change during an outbreak.

Here, real-time critical information denotes the information that is time-variant and critical to disease spreading. It includes the health condition of individuals, external environment, and social structure, etc. Real-time critical information based analysis denotes accurate characterization of disease epidemiology, prediction of trends, and evaluation of the projected impacts of different measures.

Unfortunately, real-time critical information collection and analysis remain difficult; in the past, scientists and policy-makers have often failed to make and adjust policy in real time [3].

2. Shelter System

Our Shelter (**s**martp**h**one bridg**e**d socia**l**ized body ne**t**works for epid**e**mic cont**r**ol) system has two functionality blocks; one is for critical information collection, the other is for critical information analysis.

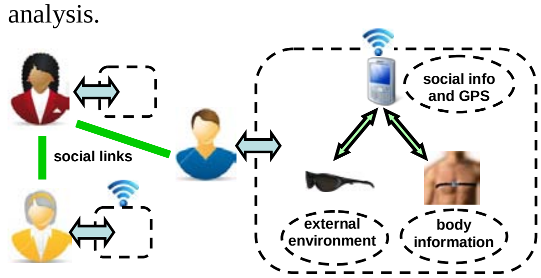


Figure 1. Critical information collection in Shelter

As illustrated in Figure 1, the information collection component in Shelter system is designed by ***seamlessly fusing body mounted sensor networks and smarphone systems***. It works as follows:

1. Smartphones collect social information. Such information includes, but not limited to, physical contact information (e.g., with whom and how long) and movement pattern.
2. Body mounted sensor networks collect body information (e.g., body temperature, blood pressure, heart rate, blood oxygen saturation) and external environment information (e.g., temperature, humility, barometric pressure, pollution level).
3. Smartphones also assist sensing body behavior or gesture information (e.g., jogging and running) and external environment information (e.g., by barometer).
4. Smartphones act as the local control center of body mounted sensor networks.
5. All sensed information will be sent to smarphones for possible preprocessing and then be uploaded to the control center.

We use smartphones as important devices (solution) to fuse the social information collection and information collection by body sensor networks because:

1. Mobile handsets, including smartphones, have become increasingly ubiquitous and feature-rich in the last ten years. We envision that this

trend will continue.

2. Communication features such as Bluetooth and Wi-Fi enable smartphone users to directly connect with each other without requiring signal coverage from base stations. The costs to build additional infrastructure for mobile social networking is reduced, which makes our system more feasible.
3. The accelerometers, GPS systems and other sensors embedded in smartphones provide users with the capability to detect surrounding people and objects with increasing accuracy and completeness (e.g., motion detection by accelerometers and people's localization by GPS systems).
4. Smartphones are becoming affordable to most of people. The cost for purchasing a smartphone has gradually decreased and will continue to decrease.

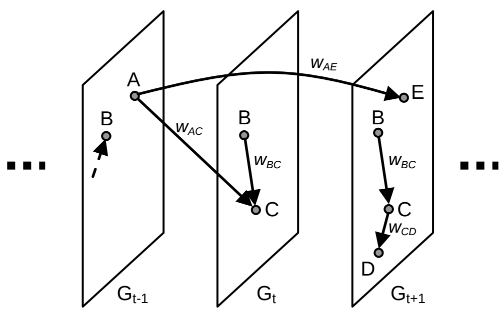


Figure 2. Critical Networks for analysis in Shelter

To capture the real-time critical information, we introduce a new concept: ***Critical Networks***. We use the example shown in Fig. 2 to illustrate the concept. The elements of critical networks are nodes, arcs, and a series of planes. A node represents a user of our Shelter system. Each node can be white, gray, or black. White nodes denote people who have never been infected and are healthy; grey nodes denote people who were infected but now have already recovered; black nodes denotes people who are infected and have not recovered yet. An arc between two nodes represents the infection relationship; the node pointed by the arc is infected by the one that originates the arc. We also define weights (using G-causality) on each arc to indicate how much in terms of possibility the node at the head of arc can affect the node at the tail of the arc. A plane represents the above information during one time period unit, which can be considered a *snapshot* of the critical networks.

In Fig. 2, snapshot G_{t-1} indicates the states and relationships of different nodes during time (day) *t-1*: node *A* and *B* are both black (sick) nodes. In snapshot G_t, node *A* recovered (become a white node and not shown in Fig. 2 for simplicity) and node *B* is still sick. We note that there appears a newly infected node *C*. *C* is originally a white node. It contacted *A* in G_{t-1}, but not got sick in G_{t-1} due to the latency and incubation nature of epidemic disease. In G_t, *C* also contact *B*. The infection of *C* can be the result of contact with A and *B* both or one of them. Such possibility is reflected by weights w_{AC} and w_{BC}. In snapshot G_{t+1}, *C* infects white node *D* with weight w_{CD}. Meanwhile, node *E* is sick due to its contact with *A* two time period units (days) early.

The critical networks enable recording the real-time critical information. It not only records the isolated snapshot information, but also records the causal relationship between different snapshots (time series). We let each node associate with a vector, which contains the states, i.e., sensed information by body mounted sensor networks, of the corresponding person at different timestamps. The arcs and their weights are calculated by integrating these states and social information such as contact and movement patterns, etc.

Based on critical networks, we can do prediction and implement strategies for epidemic control in the control center. Decisions or alerts or behavior guides that are generated by the control center will be pushed back to smartphone users.

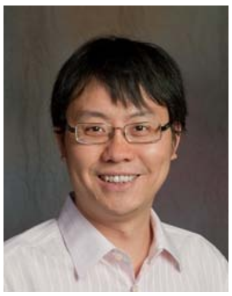
4. Final Remarks

The proposed Shelter system, by filling the dual gaps between social information collection and body including external environment information collection, and between critical information collection and critical information analysis, provides a promising solution for real time epidemic control.

Meanwhile, it raises several research challenges that need innovative solutions: 1) how to effectively and efficiently collect accurate social information, 2) how to design power saving schemes at no much cost of data collection, 3) how to efficiently and accurately present, interpret, and use of collected data, and 4) how to design fast and robust algorithms for conducting prediction and developing real-time strategies. These challenges guide our on-going and future research directions.

References

[1] Edward H. Kaplan, David L. Craft, and Lawrence M. Wein, Emergency response to a smallpox attack: The case for mass vaccination. Proceedings of the National Academy of Sciences (PNAS), 99(16):10935-10940, August 2002.
[2] Zhigang Shao, Zhijie Tan, Xianwu Zou, Zhunzhi Jin, Epidemics with pathogen mutation on small-world networks. Physics A: Statistical Mechanics and its Applications, 363(2):561–566, May 2006.
[3] Harvey V. Fineberg and Mary Elizabeth Wilson, Epidemic science in real time. Science, 324(5930):987, May 2009.

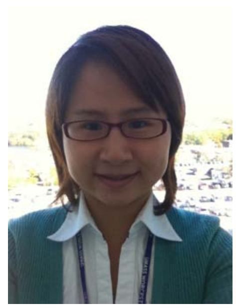


Xiaole Bai is an assistant professor in Computer and Information Science Department at University of Massachusetts Dartmouth.

He received his B.S.degree in 1999 at Southeast University, China, and the M.S. degree in 2003 at Networking Laboratory, Helsinki University of Technology, Finland. He then joined The Ohio State University in 2004 and received his Ph.D. degree in 2009 from Department of Computer Science and Engineering. His research interests include network science and engineering, cyber space security, and distributed computing.

Honggang Wang is an assistant professor in Electrical and Computer Engineering Department at University of Massachusetts Dartmouth.

He received his B.E. and M.S. degrees in Electrical Engineering, Computer Science and Engineering from Southwest Jiaotong University, China, in 1996 and 2001, respectively. He received his Ph.D. in Computer Engineering at University of Nebraska-Lincoln in 2009. His research interests include wireless sensor networks, multimedia communication, network and information security, biomedical computing, and pattern recognition. He a co-recipient of the Best Paper Award of 2008 IEEE Wireless Communications and Networking Conference (WCNC). He serves as an Associate Editor of Wiley's Security and Communication Networks (SCN) Journal, the TPC member for IEEE Globecom 2010, IEEE ICC 2011 and a Co-chair of IWCMC 2010 multimedia over wireless symposium.



Hua Fang is an Assistant Professor in Division of Biostatistics and Health Services Research, Department of Quantitative Health Science. She received her Ph.D degree in Ohio University, 2006, specialized in computational statistics, pattern recognition, research design, statistical modeling and analyses in clinical and translational research. Her research involves developing innovative methods and applying emerging robust techniques to enable or improve health studies. She won Layman Awards for missing data modeling and growth trajectory pattern recognition in 2008. She won a paper award at the 2006 Joint Research Conference on Statistics in quality industry and technology. She has been the investigator, biostatistician, leading methodologist of NIH-funded grants, currently a NIH CTSA CER Key Function Committee member in Methods workgroup. She has been a statistical consultant in health, medicine, and bio-engineering areas for multiple years.

Frequency-Aware Adaptive Indexing for P2P VoD Service with Frequent Seeks

Hongfang Guo,Zhengzhou Univesity, China
Jiangchuan Liu, Simon Fraser University, Canada
Yusong Lin and Zongmin Wang, Zhengzhou Univeristy, China
zzughf@ha.edu.cn, jcliu@cs.sfu.ca,{lys, zmwang}@zzu.edu.cn

1. Introduction

Given the frequent and complicated users' interactive operations, providing real-time Video-on-Demand (VoD) service is more challenging than live streaming, particularly for peer-to-peer VoD [1][2]. We have examined a large collection of traces from a real VoD system deployment, and have found that over 48% of user requests are indeed seeking operations. It makes the streaming non-sequential, and in turn substantially reduces the stability of a P2P overlay. The problem is further complicated by the fact that User Request Access Pattern (URAP) follows a non-uniform distribution, because users often skip boring scenes and just playback scenes of interest, even multiple times for some highlighted scenes, which has also been validated by our PPVA traces, see figure 1. Recent studies have suggested that URAP exhibits a lognormal distribution [2].

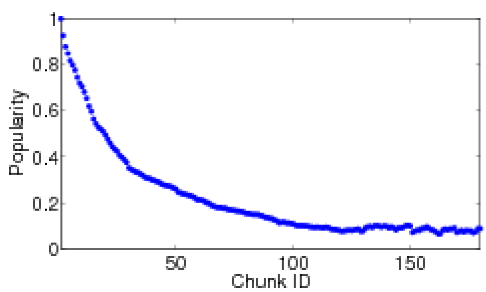


Figure1. Chunk popularity distribution

For VoD users, a small start-up latency and good streaming continuity are expected. An efficient chunk discovery upon seekings thus becomes a critical issue and an important research area in P2P VoD design.

There are many recent proposals, ranging from special overlay organization, caching, to pre-fetching and replication [3, 4, 5]. Most of them however focus only on uniform URAP, which does not reflect the real statistics. Moreover, recent studies reveal that over 80% seeking requests are of short distances [6], whose potentials and impacts have yet to be explored.

To address the above challenges, we develop a novel structure for indexing chunks for P2P VoD service with user's real-time request behaviors.

We further develop an adaptive pre-fetching policy that explores the knowledge available from the frequency-aware indexing overlay.

2. Frequency-Aware Indexing Overlay

Given the non-uniform URAP, our objective is to develop an efficient and distributed indexing structure for discovering the chunks in a peer-to-peer overlay. The indexing structure should also accommodate overlay dynamics, i.e., be self-adaptive, and should work well for unpredictable URAP, because collecting user's history behaviors is not an easy task due to the large volume of systems and the real time factor. We address these challenges by a novel indexing structure that extends a *splay tree* [7].

A basic splay tree is a self-optimizing binary search tree with the additional property that recently accessed nodes are quick to be accessed again. Once the node is accessed, it will be moved to the root of the tree through a sequence of rotation operations. In order to do this, it first performs a standard binary tree search for the node of interest, then use tree rotations to bring the node to the top. The operation of rotations is called *splaying*.

We utilize a splay tree to index chunks for the following reasons. In the VoD service, URAP is non-uniformly distributed and a splay tree can benefit VoD applications in this scenario. It ensures the chunk with high access-rate is likely to be accessed easily. Moreover, the index topology is adaptive to the user request behavior. The chunk popularity counting and extra control overhead for getting the global popularity information are not the factors affecting performance of VoD service anymore, because all these information can be extracted from the self-adaptive overlay. Finally, it is easy to implement and maintain.

Though the self-adapting nature of splay tree is attractive, applying it in a streaming overlay with interactive peers is not straightforward. First, few peers request low access-frequency chunks. Yet, once a chunk is accessed, the index node associated with this chunk should be moved to the root of the indexing tree, but due to its low popularity, the node will be splayed to the bottom again over time. This phenomenon produces unwanted operation expense, but no benefit for QoE. Second, chunks of similar popularity are requested equally, which leads to extra splaying operations frequently. Worsely, these splaying operations may not benefit the performance.

In order to address the limitations of splay tree for VoD application, we present that when a chunk is accessed at the first time, it will be inserted but will not be moved to the root directly. We limit the count of rotation operations to *count_s*, when a chunk is accessed and each rotation can make the index node associated with this chunk closer to the root with h hops, where $h=1$ *or* 2; If the chunk is of high *user* access-rate, eventually it will be moved to the root or a position close to the root after $\lceil \log C / (count\_s \times h) \rceil$ requests, where C is the number of chunks. Obviously, it filters out popular chunks and non-popular chunks self-adaptively.

Intuitively, the locations of chunks with close user request access-rate are close in indexing overlay. We consider all these chunks as a single logical chunk. Requesting any chunk of these chunks will not trigger a rotation operations and reconstruction of indexing overlay.

3. Indexing-Assisted Pre-Fetching Policy

When the buffer size of peer is limited, it is obvious that a pre-fetching strategy is a good choice to improve QoE of users. The proposed indexing overlay gives a clue about which chunk is to be pre-fetched.

The frequency-aware adaptive indexing overlay is aware of user's real time request behaviors, and it reflects the chunks' access-rate correctly without the collecting the global information of user's history behavior. Based on the knowledge about chunk access-rate extracted from the indexing overlay, we propose a simple and effective pre-fetching mechanism. Each peer will pre-fetch the chunks with high access-rate whose location is lies in the layers between the root and the current peer. If the peer has stored the some chunks among these layers, it does not need to pre-fetch them again.

4. Conclusions

We propose a novel P2P-based frequency-aware adaptive index overlay which constructs and reconstructs according to arbitrary user behavior adaptively. The advantages of this adaptive index overlay are twofold: 1) it not only ensures the quick discovery of chunks under complicated non-uniform chunk access pattern, 2) but also facilitates an efficient pre-fetching policy.

Our preliminary simulation results demonstrate its superiority for VoD application [1]: it improves the responsiveness and success rate of seeking operations, particularly for non-uniform and unpredictable user request, which outperforms other existing solutions significantly.

Acknowledgment

This research was supported by 863 Program #2008AA01A315, Henan Province Informatization Project #2008xxh001 and Henan Science and Technology Innovative Talents Program.

References

[1] H. Guo, J. Liu, and Z. Wang, Frequency-Aware Indexing for Peer-to-Peer On-Demand Video Streaming, in Proc. of IEEE ICC, 2010.

[2] A. Brampton, A. MacQuire, I. A. Rai, N. J. P. Race, L. Mathy, and M. Fry, " Characterizing User Interactivity for Sports Video-on-Demand," in Proc. of ACM NOSSDAV, 2007.

[3] X. Qiu, C. Wu, X. Lin, and F. C. Lau, " InstantLeap: Fast Neighbor Discovery in P2P VoD Streaming," in Proc. of ACM NOSSDAV, 2009.

[4] B. Cheng, H. Jin, and X. Liao, "Supporting VCR Functions in P2P VoD Services Using Ring-Assisted Overlays," in Proc. of IEEE ICC, 2007.

[5] Y. Huang, T. Z. J. Fu, D.-M. Chiu, J. C. S. Lui, and C. Huang," Challenges, Design and Analysis of a Large-scale P2P VoD System," in Proc. of ACM SIGCOMM, 2008.

[6] B. Cheng, X. Liu, Z. Zhang, and H. Jin, "A Measurement Study of a Peer-to-Peer Video-on-Demand System," in Proc. of the 6th International Workshop on Peer-to-Peer System (IPTPS), 2007.

[7] D. D. Sleator and R. E. Tarjan, "Self-Adjusting Binary Search Trees," Journal of the ACM, vol. 32, no. 3, pp. 652–686, July 1985.

Hongfang Guo received her Master's degree from Zhengzhou University in 2007, and currently is a final year PhD student at Zhengzhou University. Her research interest is on multimedia communication,

P2P streaming, Cloud computing. Prior to this, she has conducted research at Simon Fraser University, Canada, and Network Research Center of Tsinghua University, China.

Jiangchuan Liu is currently associate professor in the school of Computing Science at Simon Fraser University, B.C., Canada, From 2003 to 2004, he was an Assistant Professor in the Departmen of Computer Science and Engineering at The Chinese University of Hong Kong. He was also a Microsoft Research Fellow, and worked at Microsoft Research Asia in the summers of 2000, 2001, 2002, 2007. He received his Bachelor of Engineering (Cum Laude, Computer Science) from Tsinghua University, China, in 1999, and PhD (Computer Science) from The Hong Kong University of Science and Technology (HKUST) in 2003. He is an associate editor of IEEE Transactions on Multimedia.

Yusong Lin received his B.S. degree in Computer Science from Nanjing University in 1993 and M.S. degree from Zhengzhou University in 1998, and Ph.D. degree in Communication and Information System in National Digital Switching Research Center of China. He is currently an Associate Professor at Zhengzhou University and Associate Director at Henan Provincial Education and Research Network Center. His research is on large-scale group communication and web technologies.

Zongmin Wang is currently professor of Zhengzhou University, He received the Technology Ph.D degree from Tsinghua University in 1996. He conducted Post-Doctoral Research at Hong Kong University in 1997. His research interests include Intelligent Network Management, Multimedia Application and Virtual reality technology

LCHQ: Low Complexity with High Quality Video Codec for the Mobile Multimedia Communications

Xingang Liu, Yonsei University, Seoul, South Korea
Jianhua Ma, Hosei University, Tokyo, Japan
hanksliu@yonsei.ac.kr, jianhua@hosei.ac.jp

1. Introduction

With the fast developments of the technologies for multimedia communications and networking, the mobile video services have become very popular nowadays. As shown in Fig.1, numerous mobile video communication services have been put into use in our life, for instance, the mobile video broadcasting and cell-phone video chatting. It takes much convinces for us to know the world and communicate each other in anywhere and at anytime.

Figure1. Mobile devices in the networks

Because of the improvements of video compression techniques, the quality of the received video signal is guaranteed. Also, the communication speed is faster than before since the updated transmission technologies. The current core problems considered by the customers are the coding speed (CS) and utilization life (UL) of the mobile devices when they enjoy the high quality video programs and communications. There are so many factors which can affect the CS and UL of the mobile terminals, for instance, device quality, networking environment and software codec. Among them, the software codec is paid attention since it is more important and easier to be improved to speed-up the CS and extend the UL of the mobile devices. Recent years, the latest video coding standard, H.264/AVC [1], was implemented as the popular video compression codec because of its improved coding efficiency. It employs various techniques, such as de-blocking filter, multiple reference frames, various block size prediction and so on, to increase the coding performance. It can achieve the same quality with only half bitrate of the previous video coding standards. And it has been the dominant role and developing trend of the video signal compression. However, the complexity load of H.264/AVC encoder is really heavy due to the utilizations of the new coding tools, especially, the mode prediction and the motion estimation. With the statistical analysis, more than 90% and 80% of the entire encoding time are cost by them for Intra and Inter frames, respectively. Therefore, how to reduce the complexity has became a hot research topic for the multimedia communications and consumer electronics.

2. Research on Complexity Reduction

In fact, the state-of-the-art video compression standards, H.264/AVC and MVC, relax the constraints of the computational complexity and achieve high improvement on the video quality with the cost of huge computation. For instance, Fig.2 shows the multiple mode candidates for H.264/AVC mode prediction and motion estimation operations. To achieve the best Rate-to-Distortion Optimization (RDO), the RD Cost (RDCost) of each mode candidate should be calculated and the one which has the minimum RDCost value will be selected as the final optimal mode to guarantee the coding efficiency. According to H.264/AVC reference software [2], the RDCost computing loads are more than 4 millions units (4x4 block size is regarded as one computing unit) for just one second HDTV program which is so huge computational complexity that it effects the applications of H.264/AVC especially for the real-time mobile communications. Therefore, the main research point for the codec complexity reduction lies on how to reduce the redundancy of the mode prediction for both Intra and Inter frames due to their large probability among the entire coding procedure.

The most efficient idea is to reduce the original mode candidates because the encoding complexity comes from the RDCost computing on each of them. Many algorithms which can be categorized as following have been proposed to address the huge mode prediction problem for the video encoder so that to reduce the power consumption and extend the utilization time of the mobile devices.

- Pre-determine the prediction size.
  The multiple block sizes are supported for both Intra and Inter mode decisions, from large to small. To find the optimal mode among the multiple block sizes mode candidates while avoiding the full search algorithm, pre-determine the approximate prediction size is necessary.
- Pre-determine the mode directions
  Various directions exist in the original mode candidates. For instance, there are 9 prediction directions for Intra4x4 of H.264 Intra prediction, and the complex RDCost computing should be implemented to each of them. Therefore, pre-determine the prediction directions are also important to further reduce the mode decision computing load.

According to the context, we refer to the low complexity with high quality (LCHQ) video codec for the multimedia communication systems, especially for the power limited mobile devices. The target of LCHQ is to guarantee high quality video signal during the mobile video communications while reducing the encoding complexity at the terminals. Base on this principle, several efficient LCHQ models have been proposed in [3] with the main functionality of reducing the original mode candidates to eliminate redundant RDCost computing to save the encoding time.

3. Trend of the related technique in the future

In general, the encoding load of the video coding standards mainly lies on the transform mode prediction and motion estimation. Therefore, optimizations of these operations are the key points of the LCHQ video codec. There are two kinds of video standards currently. One is the standard which has single transform and prediction types, such as MPEG-2, H.263 and so on, where any input video sequence goes the same procedure. Although it is simple and low cost, this kind of standard can not always guarantee the video quality due to the impertinent coding procedure. On the other hand, the video coding standard, such as H.264/AVC and MVC, provides multiple candidates for the mode prediction and motion estimation. The coding efficiency can be increased significantly because of the multiple selections. However, the huge computational complexity becomes the new problem for the encoder due to the computing for the multiple mode selections. According to these scenarios, the establishment of the future LCQH video codec should utilize the advantages of the current standards while avoiding the shortcomings of them.

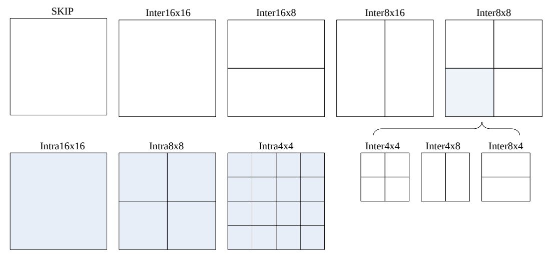


Figure 2. Multiple mode candidates in H.264/AVC encoder

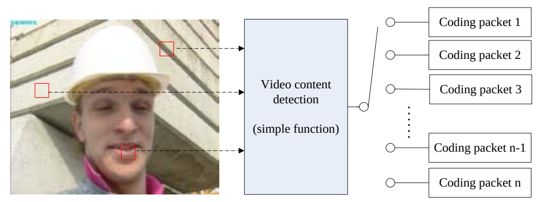


Figure 3. Illumination of LCHQ: Adaptive switches of the video coding

It is the fact that the characters among various regions of one frame are normally different, and the most suitable coding method for each region of the frame should be used to achieve the best coding efficiency. A test video frame is shown on the left of Fig.3, where the potential optimal modes and motion vectors of the three highlighted blocks are different. Instead of setting multiple modes and motion vectors as the candidates, LCHQ codec will more consider the content of the input video sequences before starting the compression. The characters, such as change trend of the video content, the relationship between current and neighboring MBs, should be measured by simple functions to pre-determine the methods of the coding operations. After that, the coding packet as shown on the right of Fig.3 is setting carefully for each regions, such as transform, mode prediction and so on. It should be noted that multiple selections for each encoding operation should be supported in the encoder database. For example, multiple transform matrixes which different sizes, transform cores and directions, and each region will select only the most suitable one for transform according to the content measurement results. Similarly, multiple prediction candidates should also be available in the standard, but instead of implementing all of the candidates to find the optimal one, only one of them is selected according to the video content measurement result to save the prediction computational load while maintaining the coding efficiency.

4. Conclusions

The increased requirements of customers need the high quality multimedia communication services in the future including the mobile communication systems. Therefore, high performance video coding standard should be employed in the mobile devices for video signal compression and decompression. However, the fixed-thinking of the tradeoff between the coding efficiency and coding complexity limits the implementation of the high efficiency coding standard. Therefore, how to solve the huge computational complexity problem becomes the critical task currently especially for the mobile terminals and embedded systems which have limited power supply and low processing ability. It is gratifying that many researchers are dedicating on this area, and the LCHQ video codec becomes the frontier for the future mobile multimedia communications. The main contribution of LCHQ video codec is to guarantee the quality of video signal during the real-time communications while reducing the codec cost to extend the energy life of the mobile devices. With the amazing increase of the mobile users, we believe that LCHQ will be an interesting and important research topic. Also, how to realize and improve the performance of LCHQ video codec will be the challenges of the future related research.

References

[1] Draft ITU-T recommendation and final draft standard of joint video dpecification, in: Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, JVT G050, 2003.

[2] Reference Software to Committee Draft Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, JVT-F100 JM10.2, 2006.

[3] D. Wu, F. Pan, K.P. Lim, S. Wu, Z.G. Li, X. Lin, S. Rahardja, C.C. Ko, "Fast intermode decision in H.264/AVC video coding," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 15, no. 6, pp. 953-958, July 2007

Xingang Liu is a BK21 research fellow in the school of electrical and electronic engineering at Yonsei University, Seoul, South Korea. He received the B.E. degree in electronics engineering from UESTC in 2000, and the M.S. and PhD degrees in the department of Information and Communication Engineering at Yeungnam University, South Korea in 2005 and 2010. During 2000-2003, he was a faculty member in the Department of Electronic Engineering of UESTC from 2000 to 2003.

Jianhua Ma is a professor at the Faculty of computer and information Sciences of Hosei University since 2000. Previously, he had 15 years' teaching/research experience at NUDT, Xidian University and the University of Aizu (Japan). His research from 1983 to 2003 covered wireless communications, data encryption, speech processing, multimedia QoS, 1-to-m HC hyper-interface, graphics ASIC, e-learning and virtual university, CSCW, multi-agents, Internet audio/video, mobile web service, P2P network, etc. Since 2003 he has been devoted to what he called Hyper World and Smart World pervaded with smart physical u-things, characterized with Ubiquitous Intelligence (ubi-Intelligence, u-Intelligence, UI) for u-Services with UbiSafe guarantee. Dr. Ma has published over 200 papers in journals and proceedings, and edited over 10 books. He is a Co-Editor-in-Chief of JUCI, JMM, JoATC, JPCC and IJUNESST. He is on the editorial boards of IJCPOL, IJDET, IJWMC, IJSH, IJSIA, IJDTA, IJCIT and IJAS, and has edited over 15 journal special issues as a Guest Editor. He is the Chair of IEEE CIS Task Force on Autonomic and Trusted Computing, and a co-founder of IEEE Task Force on Ubiquitous Intelligence and Computing.

3D Motion-Sensing Interactive HMI for Applications in Multimedia Services

Ying-Xun Lai, Wei-Guang Teng, Yueh-Min Huang, National Cheng-Kung University, Taiwan

eetaddy@gmail.com, wgteng@mail.ncku.edu.tw and huang@mail.ncku.edu.tw

1. Introduction

With the rapid development of science and technology, as well as increasing customer demand on the quality of multimedia entertainment, interactive multimedia and services become a popular topic for research. It is an interesting and challenging research topic that how to allow users to have realistic interactions with virtual multimedia services. In 2006, Nintendo introduced the new generation Wii console. The Wii Remote is a controller that allows players to interact with and manipulate items on a screen via gesture recognition. Then, motions are detected by hand-held units through screen autorotation by G-Sensor. The recent PlayStation Move and Xbox Kinect [1] provide interactive services through image identification. The above-mentioned devices emphasize on users' self-experience in the digital world, as the sensors sense users' motions and the systems reacts accordingly; however, such devices cannot achieve realistic feelings in response to occurrences in the digital world. This study aimed to design a 3D motion-sensing interactive human-machine interface (HMI) system for computers and 3D multimedia service applications. The external design connects a bending sensor on gloves, which recognizes gestures combined with a 3-axis accelerometer, achieves 3D positioning through infrared ray which improves isokinetic motion, and realizes sensing factions by spring wire and a small motor. When users select objects using the glove, the touch sensation could be realistically felt by users.

2. Related Works

HMI is an interface system for transmitting orders and commands between human and machine, enabling more complete communication between them [2]. For an interactive multimedia system, a satisfactory HMI is a necessary component. The multimedia control service is accomplished through multi-touch [3][4] and a convenient HMI system is achieved by eye movement [5][6]. 3D multimedia services have been studied more extensively in recent years. Sreeram Sreedharan et al. [7] proposed the idea of a 3D virtual world, and employed Wii Remote in the virtual environment by performing hand gestures for Yes or No gestures in order to facilitate communication. Due to the advantages of convenience and portability, the data glove [8][9], along with its sensors or other interface devices, has become a widely used interface for digital multimedia services.

3. System Implementation

The system architecture, as shown in Fig. 1, is comprised of three parts, namely, a data glove, a development board, and a computer. The glove has bending sensors and a 3-axis accelerometer, which obtains data from A/D conversion through the development board, conducts corresponding operations and normalization, and transmits the results, via Zigbee dongle, to a computer. IR (infrared LED) on the glove continuously conducts 3D positioning through an infrared camera on the computer. When the computer senses user's selection of objects, the message relayed via the Zigbee dongle of the computer to the development board, which drives the motor to draw the spring wire. Through information transmission, the bidirectional interactive motion-sensing HMI system is integrated. The descriptions of these three parts are detailed in the following.

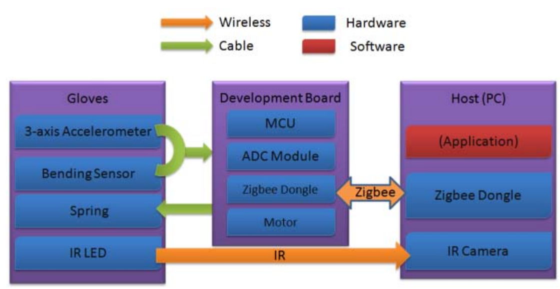


Figure 1. System architecture

■ **Data Glove.**

This study collected hand motion data through the data glove for touch sensation through the pulling force of a spring bar when the computer senses the selection of objects. The data glove, as shown in Figure 2(a), is consisted of a 3-axis accelerometer, bending sensors, spring bar, and IR LED. The 3-axis accelerometer is used to calculate the accelerating motion of the hand, determining the direction of hand motion, and attaining the angle of rotation of roll and pitch according to the gravity component theorem, as presented in Eqs.(1)(2).

$$C_{Roll} = \tan^{-1}\frac{A_z}{A_x} \text{..................}(1)$$

$$C_{pitch} = \tan^{-1}\frac{A_z}{A_x} \text{..................}(2)$$

where, A_X, A_Y, and A_Z denote the acceleration values of the three axes. The bending sensors are used for measuring the degree of a bending finger. A bending curvature sensor is a sensor that measures the bending curvature of a single direction. When the sensor bends, the resistance increases. When the bending degree increases, the resistance becomes higher. The spring bar is usually relaxed; however, when the computer triggers a digital object and sends out a message, the development board drives the motor to pull the spring tight in order that users feel the touch sensation of handling the object. IR LED is used for sending IR to an IR camera of the computer, and computing the 3D shift for modifying any inaccuracies of calculations of uniform motion by the 3-axis accelerometer.

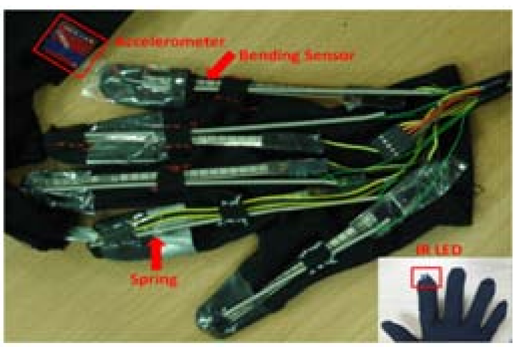


Figure 2. (a) Data glove

Figure 2. (b) Development board

■ **Development Board.**

This study used STM3210E-LK for the development platform, as shown in Figure 2(b). Its CPU is STM32F103ZE (ST Microelectronics), and is mainly for computing the hand gestures and send signals through the Zigbee communication module to the computer, and driving the motor to pull the spring. As the characteristics of bending curvature sensor's resistance may change with the degree of bending, this study connected a 15K ohm resistance in parallel, divided the voltage with the 3.3V power supply provided by STM3210E-LK development board, utilized the divided voltage to connect to STM32F103ZE microprocessor for A/D conversion, and thus, obtained the degree data of a bending finger.

Since there are errors inherent in A/D conversion, and due to different operating habits of different users, serious judgment errors may occur with the data measured by the bending sensors causing inconvenience in use. Therefore, this study used the bending motions of human fingers in clusters, where extracts of motion habits of fingers are combined with the Subtractive Clustering Method [10] of a neural network algorithm in order to determine the optimal cluster value for the boundaries of habitual motion of fingers, thus, achieving the function of self-adaptiveness.

■ **Computer**

Through the IR camera on the computer, the IR LED of the glove fingers can be accurately positioned. After the IR ray of the fingers is detected by the camera, the position of the IR can be accurately determined through appropriate procedures. However, the resolution of the screen used by users may vary with that of IR camera, and the IR resolution of 1024*768 is not completely ideal. Therefore, a scope commonly acceptable to various users must be determined, and the plane should correspond to the resolution of the output display.

In this experiment, the scope of use was (200,200)~(700,750) and assumed that the resolution of the output device is 1680*1050; and therefore, must be calculated by the following equation:

where 500 is 700-200, denoting the scope of movement on the X-axis; 550 is 750-200, denoting the scope of movement on the Y-axis. This scope may vary according to users' habits and can be adjusted. However, the adjustment should be limited in order to ensure comfort during use. X and Y are the values after conversion; Xori and Yori are the values received by the original IR. According to the above analysis, the results are summarized as follows. Suppose the comfortable scope for users is (UX1,UY1)~(UX2,UY2) and the screen resolution is W*H. And UX is defined.

Thus X and Y are calculated. Let UX = | UX2 – UX1 | and UY = | UY2 – UY1 |.

$$X=(UX2 - X_{ori}) * W/UX - W/2 \quad (3)$$

$$Y=(UY2 - Y_{ori}) * H/UY - H/2 \quad (4)$$

The deep and shallow part is calculated by the radius of the IR source point. As the hand IED is closer to the camera, the radius of light source is longer. Thus the effect can be applied to the estimation of hand motions on Z-axis to present the displacement in 3D form.

Table 1. Control commands

Function		Gesture
Select , Drag Object	Left Click	
	Drag	
Enhance Select	Right Click	
Ring Switch Windows	Menu,Confirm	
	Switch Left	
	Switch Right	
Drag Windows	Drag	
Enhance Zoom	Zoom In	
	Zoom Out	
Annotate	Write	
	Erase All	
Cube Desktop	Switch Desktop	
Mouse Control	Up	
	Down	
	Left	
	Right	
Scrolling	Scroll Up	
	Scroll Down	
Maximize Window	Slide	
Minimize Window	Slide	
Close Window	Slide	

4. Results and Conclusions

The results are as shown in Figure 3. Users can use this 3D interactive HMI for controlling the computer's operating system and achieve 3D multimedia manipulation. The control of different commands and actions by different hand gestures is shown in Table 1. This study proposed a 3D motion-sensing interactive HMI, and realized the manipulation and control of computer multimedia control system services. In the future, face and body sensing devices can be incorporated into the 3D multimedia system control services.

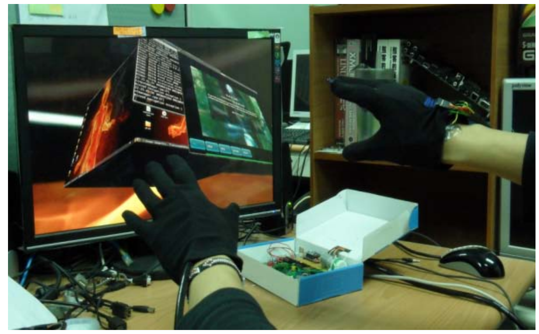


Figure 3. Achievements

Acknowledgment

This research was supported by the Sustainable Growth Project (S) (98WFA0900277) of Department of Engineering and Application Science, National Science Council (NSC), Taiwan.

References

[1] Introducing Kinect for Xbox 360 [online] http://www.xbox.com/en-US/kinect

[2] Coleman B.,"Using Sensor Inputs to Affect Virtual and Real Environments", IEEE Pervasive Computing, Volume: 8 , Issue: 3, Page(s): 16 – 23, 2009.

[3] Yu, Lingyun; Svetachov, Pjotr; Isenberg, Petra; Everts, Maarten H.; Isenberg, Tobias;"FI3D: Direct-Touch Interaction for the Exploration of 3D Scientific Visualization Spaces", IEEE Transactions on Visualization and Computer Graphics, Volume: 16, Issue: 6, Page(s): 1613 - 1622, 2010.

[4] Xing Jian-guo, Wang Wen-long, Zhao Wen-min, Huang Jing, "A Novel Multi-Touch Human-Computer-Interface Based on Binocular Stereo Vision", 2009 International Symposium on Intelligent Ubiquitous Computing and Education, Page(s): 319 – 323, 2009.

[5]Xiaoxiang Zheng, Xin Li, Jun Liu, Weidong Chen, Yaoyao Hao, "Vision-Based Multimodal Human Computer Interface Based on Parallel Tracking of Eye and Hand Motion", ICME International Conference on Complex Medical Engineering, Page(s): 1 – 5, 2009

[6] Sandnes FE, Huang YP, Huang YM, "Near Eyes-Free Chauffeur Computer Interaction with Chording and Visual Text Mnemonics" JOURNAL OF UNIVERSAL COMPUTER SCIENCE Volume: 16 Issue: 10 Pages: 1311-1326, 2010.

[7] Sreeram Sreedharan, Edmund S. Zurita, Beryl Plimmer, 3D input for 3D worlds, Proceeding OZCHI '07 Proceedings of the 19th Australasian

conference on Computer-Human Interaction: Entertaining User Interfaces, Adelaide, November, 2007.
[8]Qing Chen, Cordea M.D., Petriu E.M., Whalen T.E., Rudas I.J., Varkonyi-Koczy A., "Hand-Gesture and Facial-Expression Human-Computer Interfaces for Intelligent Space Applications", IEEE International Workshop on Medical Measurements and Applications, Page(s): 1 – 6, 2008.
[9] Zaletelj J., Perhavc J., Tasic J.F., "Vision-based human-computer interface using hand gestures", Eighth International Workshop on Image Analysis for Multimedia Interactive Services, Page(s): 41 – 41, 2007
[10] Deng Ying, Yang Shuangyuan, Liu Han, "A Subtractive Based Subspace Clustering Algorithm on High Dimensional Data", 2009 1st International Conference on Information Science and Engineering, Page(s): 766 – 769, 2009.

Ying-Xun Lai received the M.S. degree in department of electrical engineering from the National Sun Yat-sen University, Kaohsiung, Taiwan, in 2008. He is currently working toward the Ph.D. degree in engineering science from the National Cheng Kung University, Tainan, Taiwan. His main research interests are in embedded system and home digital multimedia services.

Wei-Guang Teng received the BS and PhD degrees in electrical engineering from the National Taiwan Univer sity, Taipei, Taiwan, in 1998 and 2004, respectively. He is currently an assistant professor in the Department of Engineering Science, National Cheng Kung University, Tainan, Taiwan. He was a senior researcher at the Groundhog Technologies from 2004 to 2005. His research interests include data mining, multimedia networking, and mobile computing.

Yueh-Min Huang is a Distinguished Professor in the Department of Engineering Science, National Cheng-Kung University, Taiwan, R.O.C. His research interests include wireless networks, artificial intelligence, and e-Learning. He received his MS and Ph.D. degrees in Electrical Engineering from the University of Arizona in 1988 and 1991 respectively. He has co-authored 2 books and has published over 140 refereed journal research papers. Dr. Huang has received many research awards, such as the Best Paper Award of 2007 IEA/AIE Conference, Best Paper Award of the Computer Society of the Republic of China in 2003, the Awards of Acer Long-Term Prize in 1996, 1998, and 1999, Excellent Research Awards of National Microcomputer and Communication Contests in 2006. Dr. Huang has been invited to give talks or served frequently in the program committee at national and international conferences. Dr. Huang is in the editorial board of the Journal of Wireless Communications and Mobile Computing, Journal of Security and Communication Networks and International Journal of Communication Systems. Huang is a senior member of the IEEE as well as IEEE Communication, Computer, and Circuits and Systems Societies.